(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization International Bureau



(43) International Publication Date 25 May 2001 (25.05.2001)

(51) International Patent Classification?:

PCT

C07K 14/00

(10) International Publication Number WO 01/36473 A2

(21)	International Ap	plication Number: PCT/US00/3	1581
(22)	International Fi	ling Date: 16 November 2000 (16.11.2	()(X))
(25)	Filing Language	: En	glish
(26)	Publication Language: Engli		glish
(30)	Priority Data:		
	60/165,838	16 November 1999 (16.11.1999)	US
	60/166,071	17 November 1999 (17.11.1999)	US
	60/166,678	19 November 1999 (19.11.1999)	US
	60/173,396	28 December 1999 (28.12.1999)	US
	60/184,129	22 February 2000 (22.02.2000)	US
	60/185,421	28 February 2000 (28.02.2000)	US
	60/185,554	28 February 2000 (28.02.2000)	US
	60/186,530	2 March 2000 (02:03:2000)	US
	60/186,811	3 March 2000 (03.03.2000)	US
	60/188,114	9 March 2000 (09.03.2000)	US
	60/190,310	17 March 2000 (17.03.2000)	US
	60/190,800	21 March 2000 (21.03.2000)	US
	60/198,568	20 April 2000 (20.04.2000)	US
	60/201,190	2 May 2000 (02.05.2000)	US
	60/203,111	8 May 2000 (08.05.2000)	US

Ronald, R. [US/US]: 917 Lane Boulevard. Kalamazoo MI 49001 (US). LIND, Peter [SE/SE]. Borjegatan 31C S-751 29 Uppsala (SE). SLIGHTOM, Jerry [US/US] 3305 Lorraine Avenue. Kalamazoo. MI 49008 (US) SCHELLIN, Kathleen, A. [US/US]; 361 Jason Court Portage. MI 49024 (US). KAYTES, Paul, S. [US/US] 3506-D Shadow Bend Drive, Kalamazoo. MI 49004 (US). BANNIGAN. Christopher, M. [US/US]: 2411 Kopka Court. Bay City, MI 48708 (US). RUFF, Valerie [US/US]. 7301 Sandpiper Street, Portage. MI 49024 (US) SEJLITZ, Torsten [SI/SE]: Sankt Eriksgatan 59, S-112 34 Stockholm (SE). HUFF, Rita, M. [US/US]: 3627 B Avenue, West, Plainwell, MI 49080 (US).

- (74) Agents: DELUCA, Mark et al.: Woodcock Washburn Kurtz Mackiewicz & Norris LLP, One Liberty Place, 46th Floor, Philadelphia, PA 19103 (US).
- (81) Designated States inationals: A.E. A.G. A.L. A.M. A.T. A.U. A.Z. B.A. B.B. B.G. B.R. B.Y. B.Z. C.A. C.H. C.N. C.R. C.U. C.Z. D.E. D.K. D.M. D.Z. E.E. E.S. F.L. G.B. G.D. G.E. G.H. G.M. H.R. H.U. H.D. H., H.N. IS, J.P. K.E. K.G. K.P. K.R. K.Z. L.C. L.K. L.R. S. T.T. L.U. L.V. M.A. M.D. M.G. M.K. M.N. M.W. M.X. M.Z. N.O. N.Z. P.L. P.T. R.O. R.U. S.D. S.E. S.G. S.L. S.K. S.L. T.J. T.M. T.R. T.T. T.Z. U.A. U.G. U.S. U.Z. V.N. Y.U. Z.A. Z.W.
- (84) Designated States tregionals: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW). Eurisian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM). Eurispan patent (AT, BE, CH, CY, DE, DK, ES, FL, FR, GB, GE, IE, TJ, LU, MC, NL, PT, SE, TR). OAPI patent (BE, B). CL, CG, CL, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

Without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette

(71) Applicant (for all designated States except US): PHAR-MACIA & UPJOHN COMPANY [US/US]: 301 Henrietta Street, Kalamazoo, MI 49001 (US).

25 May 2000 (25.05.2000)

(72) Inventors: and

60/207,094

(75) Inventors/Applicants (for US only): VOGELI, Gabriel [US/US]: 3005 First Avenue, Scattle, WA 98121 (US). WOOD, Linda, S. [US/US]: 10193 Fox Hollow, Portage, MI 49024 (US). PARODI, Luis, A. [SUS/E]: Greygafan:24, S-115-43 Stockholm (SI). HIEBSCH,

1/36473 A

(54) Title: NOVEL G PROTEIN-COUPLED RECEPTORS

(57) Abstract: The present invention provides a gene encoding a G protein-coupled receptor termed nGPCR-x; constructs and recombinant host cells incorporating the genes; the nGPCR-x polypeptides encoded by the gene; antibodies to the nGPCR-x polypeptides; and methods of making and using all of the foregoing.

NOVEL G PROTEIN-COUPLED RECEPTORS

CROSS-REFERENCE TO RELATED APPLICATIONS

The present application claims priority of Application Serial No. 60/165,838, filed 1999 November 16; Serial No. 60/166,071, filed 1999 November 17; Serial No. 60/166,678 filed 1999 November 19; Serial No. 60/173,396, filed 1999 December 28; Serial No. 60/184,129, filed 2000 February 22; Serial No. 60/188,114, filed 2000 March 9; Serial No. 60/185,421, filed 2000 February 28; Serial No. 60/186,811, filed 2000 March 3; Serial No. 60/186,530, filed 2000 March 2; Serial No. 60/207,094, filed 2000 May 25; Serial No. 60/203,111, filed 2000 May 8; Serial No. 60/190,310, filed 2000 March 17; Serial No. 60/201,190, filed 2000 May 2; Serial No. 60/185,554, filed 2000 February 28; Serial No. 60/198,568, filed 2000 April 20; and Serial No. 60/190,800, filed 2000 March 21, each of which is hereby incorporated by reference in its entirety.

15

20

25

30

5

10

FIELD OF THE INVENTION

The present invention relates generally to the fields of genetics and cellular and molecular biology. More particularly, the invention relates to novel G protein coupled receptors, to polynucleotides that encode such novel receptors, to reagents such as antibodies, probes, primers and kits comprising such antibodies, probes, primers related to the same, and to methods which use the novel G protein coupled receptors, polynucleotides or reagents.

BACKGROUND OF THE INVENTION

The G protein-coupled receptors (GPCRs) form a vast superfamily of cell surface receptors which are characterized by an amino-terminal extracellular domain, a carboxyl-terminal intracellular domain, and a serpentine structure that passes through the cell membrane seven times. Hence, such receptors are sometimes also referred to as seven transmembrane (7TM) receptors. These seven transmembrane domains define three extracellular loops and three intracellular loops, in addition to the amino- and carboxy- terminal domains. The extracellular portions of the receptor have a role in recognizing and binding one or more extracellular binding partners

(e.g., ligands), whereas the intracellular portions have a role in recognizing and communicating with downstream molecules in the signal transduction cascade.

The G protein-coupled receptors bind a variety of ligands including calcium ions, hormones, chemokines, neuropeptides, neurotransmitters, nucleotides, lipids, odorants, and even photons, and are important in the normal (and sometimes the aberrant) function of many cell types. [See generally Strosberg, Eur. J. Biochem. 196:1-10 (1991) and Bohm et al., Biochem J. 322:1-18 (1997).] When a specific ligand binds to its corresponding receptor, the ligand typically stimulates the receptor to activate a specific heterotrimeric guanine-nucleotide-binding regulatory protein (G-protein) that is coupled to the intracellular portion of the receptor. The G protein in turn transmits a signal to an effector molecule within the cell, by either stimulating or inhibiting the activity of that effector molecule. These effector molecules include adenylate cyclase, phospholipases and ion channels. Adenylate cyclase and phospholipases are enzymes that are involved in the production of the second messenger molecules cAMP, inositol triphosphate and diacyglycerol. It is through this sequence of events that an extracellular ligand stimuli exerts intracellular changes through a G protein-coupled receptor. Each such receptor has its own characteristic primary structure, expression pattern, ligand-binding profile, and intracellular effector system.

Because of the vital role of G protein-coupled receptors in the communication between cells and their environment, such receptors are attractive targets for therapeutic intervention, for example by activating or antagonizing such receptors. For receptors having a known ligand, the identification of agonists or antagonists may be sought specifically to enhance or inhibit the action of the ligand. Some G protein-coupled receptors have roles in disease pathogenesis (e.g., certain chemokine receptors that act as HIV co-receptors may have a role in AIDS pathogenesis), and are attractive targets for therapeutic intervention even in the absence of knowledge of the natural ligand of the receptor. Other receptors are attractive targets for therapeutic intervention by virtue of their expression pattern in tissues or cell types that are themselves attractive targets for therapeutic intervention. Examples of this latter category of receptors include receptors expressed in immune cells, which can be targeted to either inhibit autoimmune responses or to enhance immune responses to fight pathogens or cancer, and receptors expressed in the brain or other neural organs and tissues, which are likely targets in the treatment of schizophrenia, depression,

5

10

15

20

25

bipolar disease, or other neurological disorders. This latter category of receptor is also useful as a marker for identifying and/or purifying (e.g., via fluorescence-activated cell sorting) cellular subtypes that express the receptor. Unfortunately, only a limited number of G protein receptors from the central nervous system (CNS) are known. Thus, a need exists for G protein-coupled receptors that have been identified and show promise as targets for therapeutic intervention in a variety of animals, including humans.

SUMMARY OF THE INVENTION

The present invention relates to an isolated nucleic acid molecule that comprises a nucleotide sequence that encodes a polypeptide comprising an amino acid sequence homologous to even numbered sequences ranging from SEQ ID NO: 2 to SEQ ID NO: 94 and SEQ ID NO: 186, or a fragment thereof. The nucleic acid molecule encodes at least a portion of nGPCR-x. In some embodiments, the nucleic acid molecule comprises a sequence that encodes a polypeptide comprising even numbered sequences ranging from SEQ ID NO: 2 to SEQ ID NO: 94 and SEQ ID NO: 186, or a fragment thereof. In some embodiments, the nucleic acid molecule comprises a sequence homologous to odd numbered sequences ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185, or a fragment thereof. In some embodiments, the nucleic acid molecule comprises a sequence selected from the group consisting of odd numbered sequences ranging from SEQ ID NO: 1 to SEQ ID NO: 1 to SEQ ID NO: 185, and fragments thereof.

According to some embodiments, the present invention provides vectors which comprise the nucleic acid molecule of the invention. In some embodiments, the vector is an expression vector.

According to some embodiments, the present invention provides host cells which comprise the vectors of the invention. In some embodiments, the host cells comprise expression vectors.

The present invention provides an isolated nucleic acid molecule comprising a nucleotide sequence complementary to at least a portion of a sequence from an odd numbered sequence ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185, said portion comprising at least 10 nucleotides.

The present invention provides a method of producing a polypeptide comprising a sequence from an even numbered sequence ranging from SEQ ID NO: 2

5

10

15

20

25

to SEQ ID NO: 94 and SEQ ID NO: 186, or a homolog or fragment thereof. The method comprising the steps of introducing a recombinant expression vector that includes a nucleotide sequence that encodes the polypeptide into a compatible host cell, growing the host cell under conditions for expression of the polypeptide and recovering the polypeptide.

The present invention provides an isolated antibody which binds to an epitope on a polypeptide comprising a sequence from an even numbered sequence ranging from SEQ ID NO: 2 to SEQ ID NO: 94 and SEQ ID NO: 186, or a homolog or fragment thereof.

The present invention provides an method of inducing an immune response in a mammal against a polypeptide comprising a sequence from an even numbered sequence ranging from SEQ ID NO: 2 to SEQ ID NO: 94 and SEQ ID NO: 186, or a homolog or fragment thereof. The method comprises administering to a mammal an amount of the polypeptide sufficient to induce said immune response.

The present invention provides a method for identifying a compound which binds nGPCR-x. The method comprises the steps of: contacting nGPCR-x with a compound and determining whether the compound binds nGPCR-x.

The present invention provides a method for identifying a compound which binds a nucleic acid molecule encoding nGPCR-x. The method comprises the steps of contacting said nucleic acid molecule encoding nGPCR-x with a compound and determining whether said compound binds said nucleic acid molecule.

The present invention provides a method for identifying a compound which modulates the activity of nGPCR-x. The method comprises the steps of contacting nGPCR-x with a compound and determining whether nGPCR-x activity has been modulated.

The present invention provides a method of identifying an animal homolog of nGPCR-x. The method comprises the steps screening a nucleic acid database of the animal with an odd numbered sequence ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185, or a portion thereof and determining whether a portion of said library or database is homologous to said odd numbered sequence ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185, or portion thereof.

The present invention provides a method of identifying an animal homolog of nGPCR-x. The methods comprises the steps screening a nucleic acid library of the animal with a nucleic acid molecule having an odd numbered nucleotide sequence

5

10

15

20

25

ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185, or a portion thereof; and determining whether a portion of said library of database is homologous to said odd numbered nucleotide sequence ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185, or a portion thereof.

Another aspect of the present invention relates to methods of screening a human subject to diagnose a disorder affecting the brain or genetic predisposition therefor. The methods comprise the steps of assaying nucleic acid of a human subject to determine a presence or an absence of a mutation altering an amino acid sequence, expression, or biological activity of at least one nGPCR that is expressed in the brain. The nGPCR comprise an amino acid sequence selected from the group consisting of: SEQ ID NO:74, SEQ ID NO:186, SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86, SEQ ID NO:90, and SEQ ID NO:94, and allelic variants thereof. A diagnosis of the disorder or predisposition is made from the presence or absence of the mutation. The presence of a mutation altering the amino acid sequence, expression, or biological activity of the nGPCR in the nucleic acid correlates with an increased risk of developing the disorder.

The present invention further relates to methods of screening for an nGPCR-40 or nGPCR-54 hereditary schizophrenia genotype in a human patient. The methods comprise the steps of providing a biological sample comprising nucleic acid from the patient, in which the nucleic acid includes sequences corresponding to allelles of nGPCR-40 or nGPCR-54. The presence of one or more mutations in the nGPCR-40 allelle or the nGPCR-54 allelle is detected indicative of a hereditary schizophrenia genotype.

The present invention provides kits for screening a human subject to diagnose schizophrenia or a genetic predisposition therefor. The kits include an oligonucleotide useful as a probe for identifying polymorphisms in a human nGPCR-40 gene or a human nGPCR-54 gene. The oligonucleotide comprises 6-50 nucleotides in a sequence that is identical or complementary to a sequence of a wild type human nGPCR-40 or nGPCR-54 gene sequence or nGPCR-40 or nGPCR-54 coding sequence, except for one sequence difference selected from the group consisting of a nucleotide addition, a nucleotide deletion, or nucleotide substitution. The kit also includes a media packaged with the oligonucleotide. The media contains information for identifying polymorphisms that correlate with schizophrenia or a

5

10

1.5

20

25

genetic predisposition therefor, the polymophisms being identifiable using the oligonucleotide as a probe.

The present invention further relates to methods of identifying nGPCR allelic variants that correlates with mental disorders. The methods comprise the steps of providing biological samples that comprise nucleic acid from a human patient diagnosed with a mental disorder, or from the patient's genetic progenitors or progeny, and detecting in the nucleic acid the presence of one or more mutations in an nGPCR that is expressed in the brain. The nGPCR comprises an amino acid sequence selected from the group consisting of SEQ ID NO:74, SEQ ID NO:186, SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86, SEQ ID NO:90, and SEQ ID NO:94, and allelic variants thereof. The nucleic acid includes sequences corresponding to the gene or genes encoding nGPCR. The one or more mutations detected indicate an allelic variant that correlates with a mental disorder.

The present invention further relates to purified polynucleotides comprising nucleotide sequences encoding allelles of nGPCR-40 or nGPCR-54 from a human with schizophrenia. The polynucleotide hybridizes to the complement of SEQ ID NO:83 or of SEQ ID NO:85 under the following hybridization conditions: (a) hybridization for 16 hours at 42°C in a hybridization solution comprising 50% formamide, 1% SDS, 1 M NaC1, 10% dextran sulfate and (b) washing 2 times for 30 minutes at 60°C in a wash solution comprising 0.1x SSC and 1% SDS. The polynucleotide that encodes nGPCR-40 or nGPCR-54 amino acid sequence of the human differs from SEQ ID NO:84 or SEQ ID NO:86 by at least one residue.

The present invention also provides methods for identifying a modulator of biological activity of nGPCR-40 or nGPCR-54 comprising the steps of contacting a cell that expresses nGPCR-40 or nGPCR-54 in the presence and in the absence of a putative modulator compound and measuring nGPCR-40 or nGPCR-54 biological activity in the cell. The decreased or increased nGPCR-40 or nGPCR-54 biological activity in the presence versus absence of the putative modulator is indicative of a modulator of biological activity.

The present invention further provides methods to identify compounds useful for the treatment of schizophrenia. The methods comprise the steps of contacting a composition comprising nGPCR-40 with a compound suspected of binding nGPCR-40 or contacting a composition comprising nGPCR-54 with a compound suspected of

5

10

15

20

25

PCT/US00/31581

binding nGPCR-54. The binding between nGPCR-40 and the compound suspected of binding nGPCR-40 or between nGPCR-54 and the compound suspected of binding nGPCR-54 is detected. Compounds identified as binding nGPCR-40 or nGPCR-54 are candidate compounds useful for the treatment of schizophrenia.

The present invention further provides methods for identifying a compound useful as a modulator of binding between nGPCR-40 and a binding partner of nGPCR-40 or between nGPCR-54 and a binding partner of nGPCR-54. The methods comprise the steps of contacting the binding partner and a composition comprising nGPCR-40 or nGPCR-54 in the presence and in the absence of a putative modulator compound and detecting binding between the binding partner and nGPCR-40 or nGPCR-54. Decreased or increased binding between the binding partner and nGPCR-40 or nGPCR-54 in the presence of the putative modulator, as compared to binding in the absence of the putative modulator is indicative a modulator compound useful for the treatment of schizophrenia.

Another aspect of the present invention relates to methods of purifying a G protein from a sample containing a G protein. The methods comprise the steps of contacting the sample with an nGPCR for a time sufficient to allow the G protein to form a complex with the nGPCR; isolating the complex from remaining components of the sample; maintaining the complex under conditions which result in dissociation of the G protein from the nGPCR; and isolating said G protein from the nGPCR.

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS Definitions

Various definitions are made throughout this document. Most words have the meaning that would be attributed to those words by one skilled in the art. Words specifically defined either below or elsewhere in this document have the meaning provided in the context of the present invention as a whole and as are typically understood by those skilled in the art.

"Synthesized" as used herein and understood in the art, refers to polynucleotides produced by purely chemical, as opposed to enzymatic, methods. "Wholly" synthesized DNA sequences are therefore produced entirely by chemical means, and "partially" synthesized DNAs embrace those wherein only portions of the resulting DNA were produced by chemical means.

5

10

15

20

25

By the term "region" is meant a physically contiguous portion of the primary structure of a biomolecule. In the case of proteins, a region is defined by a contiguous portion of the amino acid sequence of that protein.

The term "domain" is herein defined as referring to a structural part of a biomolecule that contributes to a known or suspected function of the biomolecule. Domains may be co-extensive with regions or portions thereof; domains may also incorporate a portion of a biomolecule that is distinct from a particular region, in addition to all or part of that region. Examples of GPCR protein domains include, but are not limited to, the extracellular (*i.e.*, N-terminal), transmembrane and cytoplasmic (*i.e.*, C-terminal) domains, which are co-extensive with like-named regions of GPCRs; each of the seven transmembrane segments of a GPCR; and each of the loop segments (both extracellular and intracellular loops) connecting adjacent transmembrane segments.

5

10

15

20

25

30

NSDOCID RWO 2136473A2 File

As used herein, the term "activity" refers to a variety of measurable indicia suggesting or revealing binding, either direct or indirect; affecting a response, i.e. having a measurable affect in response to some exposure or stimulus, including, for example, the affinity of a compound for directly binding a polypeptide or polynucleotide of the invention, or, for example, measurement of amounts of upstream or downstream proteins or other similar functions after some stimulus or event.

Unless indicated otherwise, as used herein, the abbreviation in lower case (gpcr) refers to a gene, cDNA, RNA or nucleic acid sequence, while the upper case version (GPCR) refers to a protein, polypeptide, peptide, oligopeptide, or amino acid sequence. The term "nGPCR-x" refers to any of the nGPCRs taught herein, while specific reference to a nGPCR (for example nGPCR-5) refers only to that specific nGPCR.

As used herein, the term "antibody" is meant to refer to complete, intact antibodies, and Fab, Fab', F(ab)2, and other fragments thereof. Complete, intact antibodies include monoclonal antibodies such as murine monoclonal antibodies, chimeric antibodies and humanized antibodies.

As used herein, the term "binding" means the physical or chemical interaction between two proteins or compounds or associated proteins or compounds or combinations thereof. Binding includes ionic, non-ionic, Hydrogen bonds, Van der Waals, hydrophobic interactions, etc. The physical interaction, the binding, can be

WO 01/36473

either direct or indirect, indirect being through or due to the effects of another protein or compound. Direct binding refers to interactions that do not take place through or due to the effect of another protein or compound but instead are without other substantial chemical intermediates. Binding may be detected in many different manners. As a non-limiting example, the physical binding interaction between a nGPCR-x of the invention and a compound can be detected using a labeled compound. Alternatively, functional evidence of binding can be detected using, for example, a cell transfected with and expressing a nGPCR-x of the invention. Binding of the transfected cell to a ligand of the nGPCR that was transfected into the cell provides functional evidence of binding. Other methods of detecting binding are well-known to those of skill in the art.

As used herein, the term "compound" means any identifiable chemical or molecule, including, but not limited to, small molecule, peptide, protein, sugar, nucleotide, or nucleic acid, and such compound can be natural or synthetic.

As used herein, the term "complementary" refers to Watson-Crick basepairing between nucleotide units of a nucleic acid molecule.

As used herein, the term "contacting" means bringing together, either directly or indirectly, a compound into physical proximity to a polypeptide or polynucleotide of the invention. The polypeptide or polynucleotide can be in any number of buffers, salts, solutions etc. Contacting includes, for example, placing the compound into a beaker, microtiter plate, cell culture flask, or a microarray, such as a gene chip, or the like, which contains the nucleic acid molecule, or polypeptide encoding the nGPCR or fragment thereof.

As used herein, the phrase "homologous nucleotide sequence," or "homologous amino acid sequence," or variations thereof, refers to sequences characterized by a homology, at the nucleotide level or amino acid level, of at least the specified percentage. Homologous nucleotide sequences include those sequences coding for isoforms of proteins. Such isoforms can be expressed in different tissues of the same organism as a result of, for example, alternative splicing of RNA. Alternatively, isoforms can be encoded by different genes. Homologous nucleotide sequences include nucleotide sequences encoding for a protein of a species other than humans, including, but not limited to, mammals. Homologous nucleotide sequences also include, but are not limited to, naturally occurring allelic variations and mutations of the nucleotide sequences set forth herein. A homologous nucleotide sequence does

5

10

15

20

25

not, however, include the nucleotide sequence encoding other known GPCRs. Homologous amino acid sequences include those amino acid sequences which contain conscrvative amino acid substitutions and which polypeptides have the same binding and/or activity. A homologous amino acid sequence does not, however, include the amino acid sequence encoding other known GPCRs. Percent homology can be determined by, for example, the Gap program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, Madison WI), using the default settings, which uses the algorithm of Smith and Waterman (Adv. Appl. Math., 1981, 2, 482-489, which is incorporated herein by reference in its entirety).

5

10

15

20

25

30

-5000ID <WO 0136473A2 1 >

As used herein, the term "isolated" nucleic acid molecule refers to a nucleic acid molecule (DNA or RNA) that has been removed from its native environment. Examples of isolated nucleic acid molecules include, but are not limited to, recombinant DNA molecules contained in a vector, recombinant DNA molecules maintained in a heterologous host cell, partially or substantially purified nucleic acid molecules, and synthetic DNA or RNA molecules.

As used herein, the terms "modulates" or "modifies" means an increase or decrease in the amount, quality, or effect of a particular activity or protein.

As used herein, the term "oligonucleotide" refers to a series of linked nucleotide residues which has a sufficient number of bases to be used in a polymerase chain reaction (PCR). This short sequence is based on (or designed from) a genomic or cDNA sequence and is used to amplify, confirm, or reveal the presence of an identical, similar or complementary DNA or RNA in a particular cell or tissue. Oligonucleotides comprise portions of a DNA sequence having at least about 10 nucleotides and as many as about 50 nucleotides, preferably about 15 to 30 nucleotides. They are chemically synthesized and may be used as probes.

As used herein, the term "probe" refers to nucleic acid sequences of variable length, preferably between at least about 10 and as many as about 6,000 nucleotides, depending on use. They are used in the detection of identical, similar, or complementary nucleic acid sequences. Longer length probes are usually obtained from a natural or recombinant source, are highly specific and much slower to hybridize than oligomers. They may be single- or double-stranded and carefully designed to have specificity in PCR, hybridization membrane-based, or ELISA-like technologies.

WO 01/36473

The term "preventing" refers to decreasing the probability that an organism contracts or develops an abnormal condition.

The term "treating" refers to having a therapeutic effect and at least partially alleviating or abrogating an abnormal condition in the organism.

The term "therapeutic effect" refers to the inhibition or activation factors causing or contributing to the abnormal condition. A therapeutic effect relieves to some extent one or more of the symptoms of the abnormal condition. In reference to the treatment of abnormal conditions, a therapeutic effect can refer to one or more of the following: (a) an increase in the proliferation, growth, and/or differentiation of cells; (b) inhibition (i.e., slowing or stopping) of cell death; (c) inhibition of degeneration; (d) relieving to some extent one or more of the symptoms associated with the abnormal condition; and (e) enhancing the function of the affected population of cells. Compounds demonstrating efficacy against abnormal conditions can be identified as described herein.

The term "abnormal condition" refers to a function in the cells or tissues of an organism that deviates from their normal functions in that organism. An abnormal condition can relate to cell proliferation, cell differentiation, cell signaling, or cell survival. An abnormal condition may also include obesity, diabetic complications such as retinal degeneration, and irregularities in glucose uptake and metabolism, and fatty acid uptake and metabolism.

Abnormal cell proliferative conditions include cancers such as fibrotic and mesangial disorders, abnormal angiogenesis and vasculogenesis, wound healing, psoriasis, diabetes mellitus, and inflammation.

Abnormal differentiation conditions include, but are not limited to, neurodegenerative disorders, slow wound healing rates, and slow tissue grafting healing rates. Abnormal cell signaling conditions include, but are not limited to, psychiatric disorders involving excess neurotransmitter activity.

Abnormal cell survival conditions may also relate to conditions in which programmed cell death (apoptosis) pathways are activated or abrogated. A number of protein kinases are associated with the apoptosis pathways. Aberrations in the function of any one of the protein kinases could lead to cell immortality or premature cell death.

The term "administering" relates to a method of incorporating a compound into cells or tissues of an organism. The abnormal condition can be prevented or

5

10

15

20

25

treated when the cells or tissues of the organism exist within the organism or outside of the organism. Cells existing outside the organism can be maintained or grown in cell culture dishes. For cells harbored within the organism, many techniques exist in the art to administer compounds, including (but not limited to) oral, parenteral, dermal, injection, and aerosol applications. For cells outside of the organism, multiple techniques exist in the art to administer the compounds, including (but not limited to) cell microinjection techniques, transformation techniques and carrier techniques.

5

10

15

20

25

30

ENSOCGIO «WC 0136473A2 L »

The abnormal condition can also be prevented or treated by administering a compound to a group of cells having an aberration in a signal transduction pathway to an organism. The effect of administering a compound on organism function can then be monitored. The organism is preferably a mouse, rat, rabbit, guinea pig or goat, more preferably a monkey or ape, and most preferably a human.

By "amplification" it is meant increased numbers of DNA or RNA in a cell compared with normal cells. "Amplification" as it refers to RNA can be the detectable presence of RNA in cells, since in some normal cells there is no basal expression of RNA. In other normal cells, a basal level of expression exists, therefore in these cases amplification is the detection of at least 1 to 2-fold, and preferably more, compared to the basal level.

As used herein, the phrase "stringent hybridization conditions" or "stringent conditions" refers to conditions under which a probe, primer, or oligonucleotide will hybridize to its target sequence, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength, pH and nucleic acid concentration) at which 50% of the probes complementary to the target sequence hybridize to the target sequence at equilibrium. Since the target sequences are generally present in excess, at T_m, 50% of the probes are occupied at equilibrium. Typically, stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes, primers or oligonucleotides (e.g. 10 to 50 nucleotides) and at least about 60°C for longer probes, primers or oligonucleotides.

Stringent conditions may also be achieved with the addition of destabilizing agents, such as formamide.

The amino acid sequences are presented in the amino to carboxy direction, from left to right. The amino and carboxy groups are not presented in the sequence. The nucleotide sequences are presented by single strand only, in the 5' to 3' direction, from left to right. Nucleotides and amino acids are represented in the manner recommended by the IUPAC-IUB Biochemical Nomenclature Commission or (for amino acids) by three letters code.

Polynucleotides

5

10

15

20

25

30

NSDOC 5 PW 9 2136473A2

The present invention provides purified and isolated polynucleotides (e.g., DNA sequences and RNA transcripts, both sense and complementary antisense strands, both single- and double-stranded, including splice variants thereof) that encode unknown G protein-coupled receptors heretofore termed novel GPCRs, or nGPCRs. These genes are described herein and designated herein collectively as nGPCR-x (where x is 1, 3, 4, 5, 9, 11, 12, 14, 15, 18, 16, 17, 20, 21, 22, 24, 27, 28, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41, 53, 54, 55, 56, 57, 58, 59, or 60). That is, these genes are described herein and designated herein as nGPCR-1 (also referred to as beGPCR-1), nGPCR-3 (also referred to as beGPCR-3), nGPCR-4 (also referred to as beGPCR-4), nGPCR-5 (also referred to as beGPCR-5 and TL-GPCR-5), nGPCR-9 (also referred to as beGPCR-9), nGPCR-11 (also referred to as beGPCR-11), nGPCR-12 (also referred to as beGPCR-12), nGPCR-14 (also referred to as beGPCR-14), nGPCR-15 (also referred to as beGPCR-15), nGPCR-18 (also referred to as beGPCR-18), nGPCR-16 (also referred to as beGPCR-16), nGPCR-17 (also referred to as beGPCR-17), nGPCR-20 (also referred to as beGPCR-20), nGPCR-21 (also referred to as beGPCR-21), nGPCR-22 (also referred to as beGPCR-22), nGPCR-24 (also referred to as beGPCR-24), nGPCR-27 (also referred to as beGPCR-27), nGPCR-28 (also referred to as beGPCR-28), nGPCR-31 (also referred to as beGPCR-31), nGPCR-32 (also referred to as beGPCR-32), nGPCR-33 (also referred to as beGPCR-33), nGPCR-34 (also referred to as beGPCR-34), nGPCR-35 (also referred to as beGPCR-35), nGPCR-36 (also referred to as beGPCR-36), nGPCR-37 (also referred to as beGPCR-37), nGPCR-38 (also referred to as beGPCR-38), nGPCR-40 (also referred to as beGPCR-40), nGPCR-41 (also referred to as beGPCR-41), nGPCR-53, nGPCR-54, nGPCR-55, nGPCR-56, nGPCR-57, nGPCR-58, nGPCR-59, and

nGPCR-60. Table 1 below identifies the novel gene sequence nGPCR-x designation, the SEQ ID NO: of the gene sequence, the SEQ ID NO: of the polypeptide encoded thereby, and the U.S. Provisional Application in which the gene sequence has been disclosed.

5 Table 1

nGPCR	Nucleotide Sequence (SEQ ID NO:)	Amino acid Sequence (SEQ ID NO:)	Originally filed in:	nGPCR	Nucleotide Sequence (SEQ ID NO:)	Amino acid Sequence (SEQ ID NO:)	Originally filed in:
	,	2	A	32	39	40	В
1	73	74	E	33	41	42	C
3	3	4	A	34	43	44	C
3	185	186	P	35	45	46	C
	5	6	A	36	47	48	c
5	7	8	A	37	49	50	c
5	75	76	F	38	51	52	C
9	9	10	A	- 40	53	54	c
9	77	78	G	40	83	84	J
11		12	A	41	55	56	C
	79	80	H	53	57	58	D
11	13	14	A	54	59	50	D
12	15	16	A	54	85	86	K
	17	18	A	55	61	62	D
15	17	20	A	56	63	64	D
18			B	56	87		
16	21	22	I	56	89	88 90	L M
16	81	82	В	57		66	D
17	23	24	В		65		D
20	25	26		58		68	
21	27	28	В	58	91	92	N
22	29	30	В	58	93	94	0
24	31	32	В	59	69	70	D
27	33	34	В	60	71	72	D
28	35	36	В				
31	37	38	В		L		L

I	ogend	

	A= Ser. No. 60/165,838	I= Ser. No. 60/186,530
	B= Ser. No. 60/166,071	J= Ser. No. 60/207,094
10	C= Ser. No. 60/166,678	K= Ser. No. 60/203,111
	D= Ser. No. 60/173,396	L= Ser. No. 60/190,310
	E= Ser. No. 60/184,129	M= Ser. No. 60/201,190
	F= Ser. No. 60/188,114	N= Ser. No. 60/185554
	G= Ser. No. 60/185,421	O= Ser. No. 60/190,800
15	H= Ser. No. 60/186,811	P= Ser. No. 60/198,568

When a specific nGPCR is identified (for example nGPCR-5), it is understood that only that specific nGPCR is being referred to.

As described in Example 4 below, the genes encoding nGPCR-1 (nucleic acid sequence SEQ ID NO: 1, SEQ ID NO: 73, amino acid sequence SEQ ID NO: 2, SEQ ID NO:74), nGPCR-9 (nucleic acid sequence SEQ ID NO:9, SEQ ID NO:77, amino acid sequence SEQ ID NO:10, SEQ ID NO:78), nGPCR-11 (nucleic acid sequence

SEQ ID NO:11, SEQ ID NO:79, amino acid sequence SEQ ID NO:12, SEQ ID NO:80), nGPCR-16 (nucleic acid sequence SEQ ID NO: 21, SEQ ID NO:81, amino acid sequence SEQ ID NO: 22, SEQ ID NO:82), nGPCR-40 (nucleic acid sequence SEQ ID NO:53, SEQ ID NO:83, amino acid sequence SEQ ID NO:54, SEQ ID NO:84), nGPCR-54 (nucleic acid sequence SEQ ID NO:59, SEQ ID NO:85, amino acid sequence SEQ ID NO:60, SEQ ID NO: 86), nGPCR-56 (nucleic acid sequence SEQ ID NO:63, SEQ ID NO:87, SEQ ID NO:89, amino acid sequence SEQ ID NO:64, SEQ ID NO:88, SEQ ID NO:90), nGPCR-58 (nucleic acid sequence SEQ ID NO:67, SEQ ID NO:91, SEQ ID NO:93, amino acid sequence SEQ ID NO:67, SEQ ID NO:91, SEQ ID NO:93, amino acid sequence SEQ ID NO:68, SEQ ID NO: 92, SEQ ID NO:94) and nGPCR-3 (nucleic acid sequence SEQ ID NO:3, SEQ ID NO:185, amino acid sequence SEQ ID NO:4, SEQ ID NO:186) have been detected in brain tissue indicating that these n-GPCR-x proteins are neuroreceptors.

The invention provides purified and isolated polynucleotides (e.g., cDNA, genomic DNA, synthetic DNA, RNA, or combinations thereof, whether single- or double-stranded) that comprise a nucleotide sequence encoding the amino acid sequence of the polypeptides of the invention. Such polynucleotides are useful for recombinantly expressing the receptor and also for detecting expression of the receptor in cells (e.g., using Northern hybridization and in situ hybridization assays). Such polynucleotides also are useful in the design of antisense and other molecules for the suppression of the expression of nGPCR-x in a cultured cell, a tissue, or an animal: for therapeutic purposes; or to provide a model for diseases or conditions characterized by aberrant nGPCR-x expression. Specifically excluded from the definition of polynucleotides of the invention are entire isolated, non-recombinant native chromosomes of host cells. A preferred polynucleotide has the sequence of the sequence set forth in odd numbered sequences ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185, which correspond to naturally occurring nGPCR-x sequences. It will be appreciated that numerous other polynucleotide sequences exist that also encode nGPCR-x having the sequence set forth in even numbered sequences ranging from SEQ ID NO: 2 to SEQ ID NO: 94 and SEQ ID NO: 186, due to the well-known degeneracy of the universal genetic code.

The invention also provides a purified and isolated polynucleotide comprising a nucleotide sequence that encodes a mammalian polypeptide, wherein the polynucleotide hybridizes to a polynucleotide having the sequence set forth in odd numbered sequences ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID

5

10

15

20

25

NO: 185 or the non-coding strand complementary thereto, under the following hybridization conditions:

5

10

15

20

25

30

OCID <WO 0136475A2 | 5

(a) hybridization for 16 hours at 42°C in a hybridization solution comprising 50% formamide, 1% SDS, 1 M NaCl, 10% dextran sulfate; and

(b) washing 2 times for 30 minutes each at 60°C in a wash solution comprising 0.1% SSC, 1% SDS. Polynucleotides that encode a human allelic variant are highly preferred.

The present invention relates to molecules which comprise the gene sequences that encode the nGPCRs; constructs and recombinant host cells incorporating the gene sequences; the novel GPCR polypeptides encoded by the gene sequences; antibodies to the polypeptides and homologs; kits employing the polynucleotides and polypeptides, and methods of making and using all of the foregoing. In addition, the present invention relates to homologs of the gene sequences and of the polypeptides and methods of making and using the same.

Genomic DNA of the invention comprises the protein-coding region for a polypeptide of the invention and is also intended to include allelic variants thereof. It is widely understood that, for many genes, genomic DNA is transcribed into RNA transcripts that undergo one or more splicing events wherein intron (*i.e.*, non-coding regions) of the transcripts are removed, or "spliced out." RNA transcripts that can be spliced by alternative mechanisms, and therefore be subject to removal of different RNA sequences but still encode a nGPCR-x polypeptide, are referred to in the art as splice variants which are embraced by the invention. Splice variants comprehended by the invention therefore are encoded by the same original genomic DNA sequences but arise from distinct mRNA transcripts. Allelic variants are modified forms of a wild-type gene sequence, the modification resulting from recombination during chromosomal segregation or exposure to conditions which give rise to genetic mutation. Allelic variants, like wild type genes, are naturally occurring sequences (as opposed to non-naturally occurring variants that arise from *in vitro* manipulation).

The invention also comprehends cDNA that is obtained through reverse transcription of an RNA polynucleotide encoding nGPCR-x (conventionally followed by second strand synthesis of a complementary strand to provide a double-stranded DNA).

PCT/US00/31581

Preferred DNA sequences encoding human nGPCR-x polypeptides are set out in odd numbered sequences ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185. A preferred DNA of the invention comprises a double stranded molecule along with the complementary molecule (the "non-coding strand" or "complement") having a sequence unambiguously deducible from the coding strand according to Watson-Crick base-pairing rules for DNA. Also preferred are other polynucleotides encoding the nGPCR-x polypeptide of even numbered sequences ranging from SEQ ID NO: 2 to SEQ ID NO: 94 and SEQ ID NO: 186, which differ in sequence from the polynucleotides of odd numbered sequences ranging from SEQ ID NO: 93 and SEQ ID NO: 185, by virtue of the well-known degeneracy of the universal nuclear genetic code.

5

10

15

20

25

30

BNSCTOC-4WT

The invention further embraces other species, preferably mammalian, homologs of the human nGPCR-x DNA. Species homologs, sometimes referred to as "orthologs," in general, share at least 35%, at least 40%, at least 45%, at least 50%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 99% homology with human DNA of the invention. Generally, percent sequence "homology" with respect to polynucleotides of the invention may be calculated as the percentage of nucleotide bases in the candidate sequence that are identical to nucleotides in the nGPCR-x sequence set forth in odd numbered sequences ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity.

Polynucleotides of the invention permit identification and isolation of polynucleotides encoding related nGPCR-x polypeptides, such as human allelic variants and species homologs, by well-known techniques including Southern and/or Northern hybridization, and polymerase chain reaction (PCR). Examples of related polynucleotides include human and non-human genomic sequences, including allelic variants, as well as polynucleotides encoding polypeptides homologous to nGPCR-x and structurally related polypeptides sharing one or more biological, immunological, and/or physical properties of nGPCR-x. Non-human species genes encoding proteins homologous to nGPCR-x can also be identified by Southern and/or PCR analysis and are useful in animal models for nGPCR-x disorders. Knowledge of the sequence of a human nGPCR-x DNA also makes possible through use of Southern hybridization or polymerase chain reaction (PCR) the identification of genomic DNA sequences

encoding nGPCR-x expression control regulatory sequences such as promoters, operators, enhancers, repressors, and the like. Polynucleotides of the invention are also useful in hybridization assays to detect the capacity of cells to express nGPCR-x. Polynucleotides of the invention may also provide a basis for diagnostic methods useful for identifying a genetic alteration(s) in a nGPCR-x locus that underlies a disease state or states, which information is useful both for diagnosis and for selection of therapeutic strategies.

5

10

15

20

25

30

IID <WO 0136473A2 | 5

According to the present invention, the nGPCR-x nucleotide sequences disclosed herein may be used to identify homologs of the nGPCR-x, in other animals, including but not limited to humans and other mammals, and invertebrates. Any of the nucleotide sequences disclosed herein, or any portion thereof, can be used, for example, as probes to screen databases or nucleic acid libraries, such as, for example, genomic or cDNA libraries, to identify homologs, using screening procedures well known to those skilled in the art. Accordingly, homologs having at least 50%, more preferably at least 60%, more preferably at least 90%, more preferably at least 95%, and most preferably at least 100% homology with nGPCR-x sequences can be identified.

The disclosure herein of full-length polynucleotides encoding nGPCR-x polypeptides makes readily available to the worker of ordinary skill in the art every possible fragment of the full-length polynucleotide.

One preferred embodiment of the present invention provides an isolated nucleic acid molecule comprising a sequence homologous to odd numbered sequences selected from the group consisting of SEQ ID NO:1 to SEQ ID NO:93, SEQ ID NO: 185, and fragments thereof. Another preferred embodiment provides an isolated nucleic acid molecule comprising a sequence selected from the group of odd numbered sequences consisting of SEQ ID NO:1 to SEQ ID NO: 93, SEQ ID NO: 185 and fragments thereof.

As used in the present invention, fragments of nGPCR-x-encoding polynucleotides comprise at least 10, and preferably at least 12, 14, 16, 18, 20, 25, 50, or 75 consecutive nucleotides of a polynucleotide encoding nGPCR-x. Preferably, fragment polynucleotides of the invention comprise sequences unique to the nGPCR-x-encoding polynucleotide sequence, and therefore hybridize under highly stringent or moderately stringent conditions only (*i.e.*, "specifically") to polynucleotides encoding nGPCR-x (or fragments thereof). Polynucleotide fragments of genomic sequences of

PCT/US00/31581

the invention comprise not only sequences unique to the coding region, but also include fragments of the full-length sequence derived from introns, regulatory regions, and/or other non-translated sequences. Sequences unique to polynucleotides of the invention are recognizable through sequence comparison to other known polynucleotides, and can be identified through use of alignment programs routinely utilized in the art, e.g., those made available in public sequence databases. Such sequences also are recognizable from Southern hybridization analyses to determine the number of fragments of genomic DNA to which a polynucleotide will hybridize. Polynucleotides of the invention can be labeled in a manner that permits their detection, including radioactive, fluorescent, and enzymatic labeling.

Fragment polynucleotides are particularly useful as probes for detection of full-length or fragments of nGPCR-x polynucleotides. One or more polynucleotides can be included in kits that are used to detect the presence of a polynucleotide encoding nGPCR-x, or used to detect variations in a polynucleotide sequence encoding nGPCR-x.

10

15

20

25

30

ENSERGE - W.J.

31 47 47 A

The invention also embraces DNAs encoding nGPCR-x polypeptides that hybridize under moderately stringent or high stringency conditions to the non-coding strand, or complement, of the polynucleotides set forth in odd numbered sequences ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185.

Exemplary highly stringent hybridization conditions are as follows: hybridization at 42°C in a hybridization solution comprising 50% formamide, 1% SDS, 1 M NaCl, 10% Dextran sulfate, and washing twice for 30 minutes at 60°C in a wash solution comprising 0.1 X SSC and 1% SDS. It is understood in the art that conditions of equivalent stringency can be achieved through variation of temperature and buffer, or salt concentration as described Ausubel *et al.* (Eds.), Protocols in Molecular Biology, John Wiley & Sons (1994), pp. 6.0.3 to 6.4.10. Modifications in hybridization conditions can be empirically determined or precisely calculated based on the length and the percentage of guanosine/cytosine (GC) base pairing of the probe. The hybridization conditions can be calculated as described in Sambrook, *et al.*, (Eds.), Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York (1989), pp. 9.47 to 9.51.

With the knowledge of the nucleotide sequence information disclosed in the present invention, one skilled in the art can identify and obtain nucleotide sequences

which encode nGPCR-x from different sources (i.e., different tissues or different organisms) through a variety of means well known to the skilled artisan and as disclosed by, for example, Sambrook et al., "Molecular cloning: a laboratory manual", Second Edition, Cold Spring Harbor Press, Cold Spring Harbor, NY (1989), which is incorporated herein by reference in its entirety.

For example, DNA that encodes nGPCR-x may be obtained by screening of mRNA, cDNA, or genomic DNA with oligonucleotide probes generated from the nGPCR-x gene sequence information provided herein. Probes may be labeled with a detectable group, such as a fluorescent group, a radioactive atom or a chemiluminescent group in accordance with procedures known to the skilled artisan and used in conventional hybridization assays, as described by, for example, Sambrook *et al.*

A nucleic acid molecule comprising any of the nGPCR-x nucleotide sequences described above can alternatively be synthesized by use of the polymerase chain reaction (PCR) procedure, with the PCR oligonucleotide primers produced from the nucleotide sequences provided herein. See U.S. Patent Numbers 4,683,195 to Mullis et al. and 4,683,202 to Mullis. The PCR reaction provides a method for selectively increasing the concentration of a particular nucleic acid sequence even when that sequence has not been previously purified and is present only in a single copy in a particular sample. The method can be used to amplify either single- or double-stranded DNA. The essence of the method involves the use of two oligonucleotide probes to serve as primers for the template-dependent, polymerase mediated replication of a desired nucleic acid molecule.

A wide variety of alternative cloning and *in vitro* amplification methodologies are well known to those skilled in the art. Examples of these techniques are found in, for example, Berger *et al.*, *Guide to Molecular Cloning Techniques*, Methods in Enzymology 152, Academic Press, Inc., San Diego, CA (Berger), which is incorporated herein by reference in its entirety.

Automated sequencing methods can be used to obtain or verify the nucleotide sequence of nGPCR-x. The nGPCR-x nucleotide sequences of the present invention are believed to be 100% accurate. However, as is known in the art, nucleotide sequence obtained by automated methods may contain some errors. Nucleotide sequences determined by automation are typically at least about 90%, more typically at least about 95% to at least about 99.9% identical to the actual nucleotide sequence

5

10

15

20

25

PCT/US00/31581 WO 01/36473

of a given nucleic acid molecule. The actual sequence may be more precisely determined using manual sequencing methods, which are well known in the art. An error in a sequence which results in an insertion or deletion of one or more nucleotides may result in a frame shift in translation such that the predicted amino acid sequence will differ from that which would be predicted from the actual nucleotide sequence of the nucleic acid molecule, starting at the point of the mutation.

The nucleic acid molecules of the present invention, and fragments derived therefrom, are useful for screening for restriction fragment length polymorphism (RFLP) associated with certain disorders, as well as for genetic mapping.

The polynucleotide sequence information provided by the invention makes possible large-scale expression of the encoded polypeptide by techniques well known and routinely practiced in the art.

Vectors

5

10

15

20

25

30

SNST JOIL - WIT

Another aspect of the present invention is directed to vectors, or recombinant expression vectors, comprising any of the nucleic acid molecules described above. Vectors are used herein either to amplify DNA or RNA encoding nGPCR-x and/or to express DNA which encodes nGPCR-x. Preferred vectors include, but are not limited to, plasmids, phages, cosmids, episomes, viral particles or viruses, and integratable DNA fragments (i.e., fragments integratable into the host genome by homologous recombination). Preferred viral particles include, but are not limited to, adenoviruses, baculoviruses, parvoviruses, herpesviruses, poxviruses, adeno-associated viruses, Semliki Forest viruses, vaccinia viruses, and retroviruses. Preferred expression vectors include, but are not limited to, pcDNA3 (Invitrogen) and pSVL (Pharmacia Biotech). Other expression vectors include, but are not limited to, pSPORTTM vectors, pGEM™ vectors (Promega), pPROEXvectors™ (LTI, Bethesda, MD), Bluescript™ vectors (Stratagene), pQE™ vectors (Qiagen), pSE420™ (Invitrogen), and pYES2™(Invitrogen).

Expression constructs preferably comprise GPCR-x-encoding polynucleotides operatively linked to an endogenous or exogenous expression control DNA sequence and a transcription terminator. Expression control DNA sequences include promoters, enhancers, operators, and regulatory element binding sites generally, and are typically selected based on the expression systems in which the expression construct is to be utilized. Preferred promoter and enhancer sequences are generally selected for the ability to increase gene expression, while operator sequences are generally selected

for the ability to regulate gene expression. Expression constructs of the invention may also include sequences encoding one or more selectable markers that permit identification of host cells bearing the construct. Expression constructs may also include sequences that facilitate, and preferably promote, homologous recombination in a host cell. Preferred constructs of the invention also include sequences necessary for replication in a host cell.

5

10

15

20

25

30

5300000 FWO 0136473A2 L G

Expression constructs are preferably utilized for production of an encoded protein, but may also be utilized simply to amplify a nGPCR-x-encoding polynucleotide sequence. In preferred embodiments, the vector is an expression vector wherein the polynucleotide of the invention is operatively linked to a polynucleotide comprising an expression control sequence. Autonomously replicating recombinant expression constructs such as plasmid and viral DNA vectors incorporating polynucleotides of the invention are also provided. Preferred expression vectors are replicable DNA constructs in which a DNA sequence encoding nGPCR-x is operably linked or connected to suitable control sequences capable of effecting the expression of the nGPCR-x in a suitable host. DNA regions are operably linked or connected when they are functionally related to each other. For example, a promoter is operably linked or connected to a coding sequence if it controls the transcription of the sequence. Amplification vectors do not require expression control domains, but rather need only the ability to replicate in a host, usually conferred by an origin of replication, and a selection gene to facilitate recognition of transformants. The need for control sequences in the expression vector will vary depending upon the host selected and the transformation method chosen. Generally, control sequences include a transcriptional promoter, an optional operator sequence to control transcription, a sequence encoding suitable mRNA ribosomal binding and sequences which control the termination of transcription and translation.

Preferred vectors preferably contain a promoter that is recognized by the host organism. The promoter sequences of the present invention may be prokaryotic, eukaryotic or viral. Examples of suitable prokaryotic sequences include the P_R and P_L promoters of bacteriophage lambda (The bacteriophage Lambda, Hershey, A. D., Ed., Cold Spring Harbor Press, Cold Spring Harbor, NY (1973), which is incorporated herein by reference in its entirety; Lambda II, Hendrix, R. W., Ed., Cold Spring Harbor Press, Cold Spring Harbor, NY (1980), which is incorporated herein by reference in its entirety); the trp, recA, heat shock, and lacZ promoters of *E. coli* and

PCT/US00/31581 WO 01/36473

the SV40 early promoter (Benoist et al. Nature, 1981, 290, 304-310, which is incorporated herein by reference in its entirety). Additional promoters include, but are not limited to, mouse mammary tumor virus, long terminal repeat of human immunodeficiency virus, maloney virus, cytomegalovirus immediate early promoter, Epstein Barr virus, Rous sarcoma virus, human actin, human myosin, human hemoglobin, human muscle creatine, and human metalothionein.

5

10

15

20

25

30

BNSCOOD WAD

1964-5A2 1 5

Additional regulatory sequences can also be included in preferred vectors.

Preferred examples of suitable regulatory sequences are represented by the ShineDalgarno of the replicase gene of the phage MS-2 and of the gene cII of
bacteriophage lambda. The Shine-Dalgarno sequence may be directly followed by
DNA encoding nGPCR-x and result in the expression of the mature nGPCR-x protein.

Moreover, suitable expression vectors can include an appropriate marker that allows the screening of the transformed host cells. The transformation of the selected host is carried out using any one of the various techniques well known to the expert in the art and described in Sambrook *et al.*, *supra*.

An origin of replication can also be provided either by construction of the vector to include an exogenous origin or may be provided by the host cell chromosomal replication mechanism. If the vector is integrated into the host cell chromosome, the latter may be sufficient. Alternatively, rather than using vectors which contain viral origins of replication, one skilled in the art can transform mammalian cells by the method of co-transformation with a selectable marker and nGPCR-x DNA. An example of a suitable marker is dihydrofolate reductase (DHFR) or thymidine kinase (see, U.S. Patent No. 4,399,216).

Nucleotide sequences encoding GPCR-x may be recombined with vector DNA in accordance with conventional techniques, including blunt-ended or staggered-ended termini for ligation, restriction enzyme digestion to provide appropriate termini, filling in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesiderable joining, and ligation with appropriate ligases. Techniques for such manipulation are disclosed by Sambrook et al., *supra* and are well known in the art. Methods for construction of mammalian expression vectors are disclosed in, for example, Okayama *et al.*, *Mol. Cell. Biol.*, 1983, 3, 280, Cosman *et al.*, *Mol. Immunol.*, 1986, 23, 935, Cosman *et al.*, *Nature*, 1984, 312, 768, EP-A-0367566, and WO 91/18982, each of which is incorporated herein by reference in its entirety.

Host cells

5

10

15

20

25

30

According to another aspect of the invention, host cells are provided, including prokaryotic and eukaryotic cells, comprising a polynucleotide of the invention (or vector of the invention) in a manner that permits expression of the encoded nGPCR-x polypeptide. Polynucleotides of the invention may be introduced into the host cell as part of a circular plasmid, or as linear DNA comprising an isolated protein coding region or a viral vector. Methods for introducing DNA into the host cell that are well known and routinely practiced in the art include transformation, transfection, electroporation, nuclear injection, or fusion with carriers such as liposomes, micelles, ghost cells, and protoplasts. Expression systems of the invention include bacterial, yeast, fungal, plant, insect, invertebrate, vertebrate, and mammalian cells systems.

The invention provides host cells that are transformed or transfected (stably or transiently) with polynucleotides of the invention or vectors of the invention. As stated above, such host cells are useful for amplifying the polynucleotides and also for expressing the nGPCR-x polypeptide or fragment thereof encoded by the polynucleotide.

In still another related embodiment, the invention provides a method for producing a nGPCR-x polypeptide (or fragment thereof) comprising the steps of growing a host cell of the invention in a nutrient medium and isolating the polypeptide or variant thereof from the cell or the medium. Because nGPCR-x is a seven transmembrane receptor, it will be appreciated that, for some applications, such as certain activity assays, the preferable isolation may involve isolation of cell membranes containing the polypeptide embedded therein, whereas for other applications a more complete isolation may be preferable.

According to some aspects of the present invention, transformed host cells having an expression vector comprising any of the nucleic acid molecules described above are provided. Expression of the nucleotide sequence occurs when the expression vector is introduced into an appropriate host cell. Suitable host cells for expression of the polypeptides of the invention include, but are not limited to, prokaryotes, yeast, and eukaryotes. If a prokaryotic expression vector is employed, then the appropriate host cell would be any prokaryotic cell capable of expressing the cloned sequences. Suitable prokaryotic cells include, but are not limited to, bacteria

PCT/US00/31581 WO 01/36473

of the genera Escherichia, Bacillus, Salmonella, Pseudomonas, Streptomyces, and Staphylococcus.

5

10

15

20

25

30

BNS0400 LIKW

113647347

If an eukaryotic expression vector is employed, then the appropriate host cell would be any eukaryotic cell capable of expressing the cloned sequence. Preferably, eukaryotic cells are cells of higher eukaryotes. Suitable eukaryotic cells include, but are not limited to, non-human mammalian tissue culture cells and human tissue culture cells. Preferred host cells include, but are not limited to, insect cells, HeLa cells, Chinese hamster ovary cells (CHO cells), African green monkey kidney cells (COS cells), human 293 cells, and murine 3T3 fibroblasts. Propagation of such cells in cell culture has become a routine procedure (see, Tissue Culture, Academic Press, Kruse and Patterson, eds. (1973), which is incorporated herein by reference in its entirety).

In addition, a yeast host may be employed as a host cell. Preferred yeast cells include, but are not limited to, the genera Saccharomyces, Pichia, and Kluveromyces. Preferred yeast hosts are S. cerevisiae and P. pastoris. Preferred yeast vectors can contain an origin of replication sequence from a 2T yeast plasmid, an autonomously replication sequence (ARS), a promoter region, sequences for polyadenylation, sequences for transcription termination, and a selectable marker gene. Shuttle vectors for replication in both yeast and E. coli are also included herein.

Alternatively, insect cells may be used as host cells. In a preferred embodiment, the polypeptides of the invention are expressed using a baculovirus expression system (*see*, Luckow *et al.*, *Bio/Technology*, **1988**, *6*, 47, Baculovirus Expression Vectors: A Laboratory Manual, O'Rielly *et al.* (Eds.), W.H. Freeman and Company, New York, **1992**, and U.S. Patent No. 4,879,236, each of which is incorporated herein by reference in its entirety). In addition, the MAXBACTM complete baculovirus expression system (Invitrogen) can, for example, be used for production in insect cells.

Host cells of the invention are a valuable source of immunogen for development of antibodies specifically immunoreactive with nGPCR-x. Host cells of the invention are also useful in methods for the large-scale production of nGPCR-x polypeptides wherein the cells are grown in a suitable culture medium and the desired polypeptide products are isolated from the cells, or from the medium in which the cells are grown, by purification methods known in the art, e.g., conventional chromatographic methods including immunoaffinity chromatography, receptor

affinity chromatography, hydrophobic interaction chromatography, lectin affinity chromatography, size exclusion filtration, cation or anion exchange chromatography, high pressure liquid chromatography (HPLC), reverse phase HPLC, and the like. Still other methods of purification include those methods wherein the desired protein is expressed and purified as a fusion protein having a specific tag, label, or chelating moiety that is recognized by a specific binding partner or agent. The purified protein can be cleaved to yield the desired protein, or can be left as an intact fusion protein. Cleavage of the fusion component may produce a form of the desired protein having additional amino acid residues as a result of the cleavage process.

Knowledge of nGPCR-x DNA sequences allows for modification of cells to permit, or increase, expression of endogenous nGPCR-x. Cells can be modified (e.g., by homologous recombination) to provide increased expression by replacing, in whole or in part, the naturally occurring nGPCR-x promoter with all or part of a heterologous promoter so that the cells express nGPCR-x at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to endogenous nGPCR-x encoding sequences. (See, for example, PCT International Publication No. WO 94/12650, PCT International Publication No. WO 92/20808, and PCT International Publication No. WO 91/09955.) It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene which encodes carbamoyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the nGPCR-x coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the nGPCR-x coding sequences in the cells.

Knock-outs

5

10

15

20

25

30

BNSCOCID 6WO 019,473A2 . .

The DNA sequence information provided by the present invention also makes possible the development (e.g., by homologous recombination or "knock-out" strategies; see Capecchi, Science 244:1288-1292 (1989), which is incorporated herein by reference) of animals that fail to express functional nGPCR-x or that express a variant of nGPCR-x. Such animals (especially small laboratory animals such as rats, rabbits, and mice) are useful as models for studying the in vivo activities of nGPCR-x and modulators of nGPCR-x.

PCT/US00/31581 WO 01/36473

Antisense

10

15

20

25

30

5N30000 - W0

2196477A7

Also made available by the invention are anti-sense polynucleotides that recognize and hybridize to polynucleotides encoding nGPCR-x. Full-length and fragment anti-sense polynucleotides are provided. Fragment antisense molecules of the invention include (i) those that specifically recognize and hybridize to nGPCR-x RNA (as determined by sequence comparison of DNA encoding nGPCR-x to DNA encoding other known molecules). Identification of sequences unique to nGPCR-x encoding polynucleotides can be deduced through use of any publicly available sequence database, and/or through use of commercially available sequence comparison programs. After identification of the desired sequences, isolation through restriction digestion or amplification using any of the various polymerase chain reaction techniques well known in the art can be performed. Anti-sense polynucleotides are particularly relevant to regulating expression of nGPCR-x by those cells expressing nGPCR-x mRNA.

Antisense nucleic acids (preferably 10 to 30 base-pair oligonucleotides) capable of specifically binding to nGPCR-x expression control sequences or nGPCRx RNA are introduced into cells (e.g., by a viral vector or colloidal dispersion system such as a liposome). The antisense nucleic acid binds to the nGPCR-x target nucleotide sequence in the cell and prevents transcription and/or translation of the target sequence. Phosphorothioate and methylphosphonate antisense oligonucleotides are specifically contemplated for therapeutic use by the invention. The antisense oligonucleotides may be further modified by adding poly-L-lysine, transferrin polylysine, or cholesterol moieties at their 5' end. Suppression of nGPCR-x expression at either the transcriptional or translational level is useful to generate cellular or animal models for diseases/conditions characterized by aberrant nGPCR-x expression.

Antisense oligonucleotides, or fragments of odd numbered nucleotide sequences ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185 or sequences complementary or homologous thereto, derived from the nucleotide sequences of the present invention encoding nGPCR-x are useful as diagnostic tools for probing gene expression in various tissues. For example, tissue can be probed in situ with oligonucleotide probes carrying detectable groups by conventional autoradiography techniques to investigate native expression of this enzyme or pathological conditions relating thereto. Antisense oligonucleotides are preferably

directed to regulatory regions of odd numbered nucleotide sequences ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185, or mRNA corresponding thereto, including, but not limited to, the initiation codon, TATA box, enhancer sequences, and the like.

Transcription factors

5

10

15

20

25

30

NEDOCID <WO - 0136473A2 1 >

The nGPCR-x sequences taught in the present invention facilitate the design of novel transcription factors for modulating nGPCR-x expression in native cells and animals, and cells transformed or transfected with nGPCR-x polynucleotides. For example, the Cvs₂-His₂ zinc finger proteins, which bind DNA via their zinc finger domains, have been shown to be amenable to structural changes that lead to the recognition of different target sequences. These artificial zinc finger proteins recognize specific target sites with high affinity and low dissociation constants, and are able to act as gene switches to modulate gene expression. Knowledge of the particular nGPCR-x target sequence of the present invention facilitates the engineering of zinc finger proteins specific for the target sequence using known methods such as a combination of structure-based modeling and screening of phage display libraries (Segal et al., Proc. Natl. Acad. Sci. (USA) 96:2758-2763 (1999); Liu et al., Proc. Natl. Acad. Sci. (USA) 94:5525-5530 (1997); Greisman et al., Science 275:657-661 (1997); Choo et al., J. Mol. Biol. 273:525-532 (1997)). Each zinc finger domain usually recognizes three or more base pairs. Since a recognition sequence of 18 base pairs is generally sufficient in length to render it unique in any known genome, a zinc finger protein consisting of 6 tandem repeats of zinc fingers would be expected to ensure specificity for a particular sequence (Segal et al.) The artificial zinc finger repeats, designed based on nGPCR-x sequences, are fused to activation or repression domains to promote or suppress nGPCR-x expression (Liu et al.) Alternatively, the zinc finger domains can be fused to the TATA box-binding factor (TBP) with varying lengths of linker region between the zinc finger peptide and the TBP to create either transcriptional activators or repressors (Kim et al., Proc. Natl. Acad. Sci. (USA) 94:3616-3620 (1997). Such proteins and polynucleotides that encode them, have utility for modulating nGPCR-x expression in vivo in both native cells, animals and humans; and/or cells transfected with nGPCR-x-encoding sequences. The novel transcription factor can be delivered to the target cells by transfecting constructs that express the transcription factor (gene therapy), or by introducing the protein. Engineered zinc finger proteins can also be designed to bind

WO 01/36473

RNA sequences for use in therapeutics as alternatives to antisense or catalytic RNA methods (McColl *et al.*, Proc. Natl. Acad. Sci. (USA) 96:9521-9526 (1997); Wu *et al.*, Proc. Natl. Acad. Sci. (USA) 92:344-348 (1995)). The present invention contemplates methods of designing such transcription factors based on the gene sequence of the invention, as well as customized zinc finger proteins, that are useful to modulate nGPCR-x expression in cells (native or transformed) whose genetic complement includes these sequences.

Polypeptides

5

10

15

20

25

30

NESCOSS RWS

The invention also provides purified and isolated mammalian nGPCR-x polypeptides encoded by a polynucleotide of the invention. Presently preferred is a human nGPCR-x polypeptide comprising the amino acid sequence set out in even numbered sequences ranging from SEQ ID NO: 2 to SEQ ID NO: 94 and SEQ ID NO: 186 or fragments thereof comprising an epitope specific to the polypeptide. By "epitope specific to" is meant a portion of the nGPCR receptor that is recognizable by an antibody that is specific for the nGPCR, as defined in detail below.

Although the sequences provided are particular human sequences, the invention is intended to include within its scope other human allelic variants; non-human mammalian forms of nGPCR-x, and other vertebrate forms of nGPCR-x.

It will be appreciated that extracellular epitopes are particularly useful for generating and screening for antibodies and other binding compounds that bind to receptors such as nGPCR-x. Thus, in another preferred embodiment, the invention provides a purified and isolated polypeptide comprising at least one extracellular domain (e.g., the N-terminal extracellular domain or one of the three extracellular loops) of nGPCR-x. Purified and isolated polypeptides comprising the N-terminal extracellular domain of nGPCR-x are highly preferred. Also preferred is a purified and isolated polypeptide comprising a nGPCR-x fragment selected from the group consisting of the N-terminal extracellular domain of nGPCR-x, transmembrane domains of nGPCR-x, an extracellular loop connecting transmembrane domains of nGPCR-x, an intracellular loop connecting transmembrane domains of nGPCR-x, the C-terminal cytoplasmic region of nGPCR-x, and fusions thereof. Such fragments may be continuous portions of the native receptor. However, it will also be appreciated that knowledge of the nGPCR-x gene and protein sequences as provided herein permits recombining of various domains that are not contiguous in the native protein. Using a FORTRAN computer program called "tmtrest.all" [Parodi et al.,

Comput. Appl. Biosci. 5:527-535 (1994)], nGPCR-x was shown to contain transmembrane-spanning domains.

5

10

15

20

25

30

INSDUCID KWO GERAZBAR . .

The invention also embraces polypeptides that have at least 99%, at least 95%, at least 90%, at least 85%, at least 80%, at least 75%, at least 70%, at least 65%, at least 60%, at least 55% or at least 50% identity and/or homology to the preferred polypeptide of the invention. Percent amino acid sequence "identity" with respect to the preferred polypeptide of the invention is defined herein as the percentage of amino acid residues in the candidate sequence that are identical with the residues in the nGPCR-x sequence after aligning both sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Percent sequence "homology" with respect to the preferred polypeptide of the invention is defined herein as the percentage of amino acid residues in the candidate sequence that are identical with the residues in the nGPCR-x sequence after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and also considering any conservative substitutions as part of the sequence identity.

In one aspect, percent homology is calculated as the percentage of amino acid residues in the smaller of two sequences which align with identical amino acid residue in the sequence being compared, when four gaps in a length of 100 amino acids may be introduced to maximize alignment [Dayhoff, in Atlas of Protein Sequence and Structure, Vol. 5, p. 124, National Biochemical Research Foundation, Washington, D.C. (1972), incorporated herein by reference].

Polypeptides of the invention may be isolated from natural cell sources or may be chemically synthesized, but are preferably produced by recombinant procedures involving host cells of the invention. Use of mammalian host cells is expected to provide for such post-translational modifications (*e.g.*, glycosylation, truncation, lipidation, and phosphorylation) as may be needed to confer optimal biological activity on recombinant expression products of the invention. Glycosylated and non-glycosylated forms of nGPCR-x polypeptides are embraced by the invention.

The invention also embraces variant (or analog) nGPCR-x polypeptides. In one example, insertion variants are provided wherein one or more amino acid residues supplement a nGPCR-x amino acid sequence. Insertions may be located at either or both termini of the protein, or may be positioned within internal regions of the nGPCR-x amino acid sequence. Insertional variants with additional residues at either

WO 01/36473

or both termini can include, for example, fusion proteins and proteins including amino acid tags or labels.

Insertion variants include nGPCR-x polypeptides wherein one or more amino acid residues are added to a nGPCR-x acid sequence or to a biologically active fragment thereof.

Variant products of the invention also include mature nGPCR-x products, *i.e.*, nGPCR-x products wherein leader or signal sequences are removed, with additional amino terminal residues. The additional amino terminal residues may be derived from another protein, or may include one or more residues that are not identifiable as being derived from specific proteins. nGPCR-x products with an additional methionine residue at position -1 (Met⁻¹-nGPCR-x) are contemplated, as are variants with additional methionine and lysine residues at positions -2 and -1 (Met⁻²-Lys⁻¹-nGPCR-x). Variants of nGPCR-x with additional Met, Met-Lys, Lys residues (or one or more basic residues in general) are particularly useful for enhanced recombinant protein production in bacterial host cells.

The invention also embraces nGPCR-x variants having additional amino acid residues that result from use of specific expression systems. For example, use of commercially available vectors that express a desired polypeptide as part of a glutathione-S-transferase (GST) fusion product provides the desired polypeptide having an additional glycine residue at position -1 after cleavage of the GST component from the desired polypeptide. Variants that result from expression in other vector systems are also contemplated.

Insertional variants also include fusion proteins wherein the amino terminus and/or the carboxy terminus of nGPCR-x is/are fused to another polypeptide.

In another aspect, the invention provides deletion variants wherein one or more amino acid residues in a nGPCR-x polypeptide are removed. Deletions can be effected at one or both termini of the nGPCR-x polypeptide, or with removal of one or more non-terminal amino acid residues of nGPCR-x. Deletion variants, therefore, include all fragments of a nGPCR-x polypeptide.

The invention also embraces polypeptide fragments of the even numbered sequences ranging from SEQ ID NO: 2 to SEQ ID NO: 94 and SEQ ID NO: 186, wherein the fragments maintain biological (e.g., ligand binding and/or intracellular signaling) immunological properties of a nGPCR-x polypeptide.

5

10

15

20

25

In one preferred embodiment of the invention, an isolated nucleic acid molecule comprises a nucleotide sequence that encodes a polypeptide comprising an amino acid sequence homologous to even numbered sequences selected from the group consisting of: SEQ ID NO:2 to SEQ ID NO:94, SEQ ID NO: 186, and fragments thereof, wherein the nucleic acid molecule encoding at least a portion of nGPCR-x. In a more preferred embodiment, the isolated nucleic acid molecule comprises a sequence that encodes a polypeptide comprising even numbered sequences selected from the group consisting of SEQ ID NO:2 to SEQ ID NO: 94, SEQ ID NO: 186, and fragments thereof.

As used in the present invention, polypeptide fragments comprise at least 5, 10, 15, 20, 25, 30, 35, or 40 consecutive amino acids of the even numbered sequences ranging from SEQ ID NO: 2 to SEQ ID NO: 94 and SEQ ID NO: 186. Preferred polypeptide fragments display antigenic properties unique to, or specific for, human nGPCR-x and its allelic and species homologs. Fragments of the invention having the desired biological and immunological properties can be prepared by any of the methods well known and routinely practiced in the art.

In still another aspect, the invention provides substitution variants of nGPCR-x polypeptides. Substitution variants include those polypeptides wherein one or more amino acid residues of a nGPCR-x polypeptide are removed and replaced with alternative residues. In one aspect, the substitutions are conservative in nature; however, the invention embraces substitutions that are also non-conservative. Conservative substitutions for this purpose may be defined as set out in Tables 2, 3, or 4 below.

Variant polypeptides include those wherein conservative substitutions have been introduced by modification of polynucleotides encoding polypeptides of the invention. Amino acids can be classified according to physical properties and contribution to secondary and tertiary protein structure. A conservative substitution is recognized in the art as a substitution of one amino acid for another amino acid that has similar properties. Exemplary conservative substitutions are set out in Table 2 (from WO 97/09433, page 10, published March 13, 1997 (PCT/GB96/02197, filed 9/6/96), immediately below.

5

10

15

20

25

Table 2

Conservative Substitutions I

SIDE CHAIN CHARACTERISTIC Aliphatic	AMINO ACID
Non-polar	GAP
rion-polar	ILV
Polar - uncharged	CSTM
1 biai - ancharges	NO
Polar - charged	DÈ
1 Olai - Charges	KR
Aromatic	HFWY
Other	NODE
Oulci	

Alternatively, conservative amino acids can be grouped as described in Lehninger, [Biochemistry, Second Edition; Worth Publishers, Inc. NY, NY (1975), pp.71-77] as set out in Table 3, below.

Table 3
Conservative Substitutions II

SIDE CHAIN AMINO ACID CHARACTERISTIC Non-polar (hydrophobic) A. Aliphatic: ALIVP B. Aromatic: F W C. Sulfur-containing: Μ D. Borderline: Uncharged-polar A. Hydroxyl. STY B. Amides: ΝQ C C. Sulfhydryl: D. Borderline: G KRH Positively Charged (Basic):

Negatively Charged (Acidic):

As still another alternative, exemplary conservative substitutions are set out in Table 4, below.

Table 4
Conservative Substitutions III

Original Residue	Exemplary Substitution
Ala (A)	Val, Leu, Ile
Arg (R)	Lys, Gln, Asn
Asn (N)	Gln, His, Lys, Arg
Asp (D)	Glu
Cys (C)	Ser
Gln (Q)	Asn
Glu (E)	Asp
His (H)	Asn, Gln, Lys, Arg
Ile (I)	Leu, Val, Met, Ala, Phe,

DE

5

10

Leu (L)	Ile, Val, Met, Ala, Phe
Lys (K)	Arg, Gln, Asn
Met (M)	Leu, Phe, Ile
Phe (F)	Leu, Val, Ile, Ala
Pro (P)	Gly
Ser (S)	Thr
Thr (T)	Ser
Trp (W)	Tyr
Tyr (Y)	Trp, Phe, Thr, Ser
Val (V)	Ile, Leu, Met, Phe, Ala

It should be understood that the definition of polypeptides of the invention is intended to include polypeptides bearing modifications other than insertion, deletion, or substitution of amino acid residues. By way of example, the modifications may be covalent in nature, and include for example, chemical bonding with polymers, lipids, other organic, and inorganic moieties. Such derivatives may be prepared to increase circulating half-life of a polypeptide, or may be designed to improve the targeting capacity of the polypeptide for desired cells, tissues, or organs. Similarly, the invention further embraces nGPCR-x polypeptides that have been covalently modified to include one or more water-soluble polymer attachments such as polyethylene glycol, polyoxyethylene glycol, or polypropylene glycol. Variants that display ligand binding properties of native nGPCR-x and are expressed at higher levels, as well as variants that provide for constitutively active receptors, are particularly useful in assays of the invention; the variants are also useful in providing cellular, tissue and animal models of diseases/conditions characterized by aberrant nGPCR-x activity.

5

10

15

20

25

: DOCID <WO 0136473A2 | s

In a related embodiment, the present invention provides compositions comprising purified polypeptides of the invention. Preferred compositions comprise, in addition to the polypeptide of the invention, a pharmaceutically acceptable (*i.e.*, sterile and non-toxic) liquid, semisolid, or solid diluent that serves as a pharmaceutical vehicle, excipient, or medium. Any diluent known in the art may be used. Exemplary diluents include, but are not limited to, water, saline solutions, polyoxyethylene sorbitan monolaurate, magnesium stearate, methyl- and propylhydroxybenzoate, talc, alginates, starches, lactose, sucrose, dextrose, sorbitol, mannitol, glycerol, calcium phosphate, mineral oil, and cocoa butter.

Variants that display ligand binding properties of native nGPCR-x and are expressed at higher levels, as well as variants that provide for constitutively active

receptors, are particularly useful in assays of the invention; the variants are also useful in assays of the invention and in providing cellular, tissue and animal models of diseases/conditions characterized by aberrant nGPCR-x activity.

The G protein-coupled receptor functions through a specific heterotrimeric guanine-nucleotide-binding regulatory protein (G-protein) coupled to the intracellular portion of the G protein-coupled receptor molecule. Accordingly, the G protein-coupled receptor has a specific affinity to G protein. G proteins specifically bind to guanine nucleotides. Isolation of G proteins provides a means to isolate guanine nucleotides. G Proteins may be isolated using commercially available anti-G protein antibodies or isolated G protein-coupled receptors. Similarly, G proteins may be detected in a sample isolated using commercially available detectable anti-G protein antibodies or isolated G protein-coupled receptors.

According to the present invention, the isolated n-GPCR-x proteins of the present invention are useful to isolate and purify G proteins from samples such as cell lysates. Example 15 below sets forth an example of isolation of G proteins using isolated n-GPCR-x proteins. Such methodolgy may be used in place of the use of commercially available anti-G protein antibodies which are used to isolate G proteins. Moreover, G proteins may be detected using n-GPCR-x proteins in place of commercially available detectable anti-G protein antibodies. Since n-GPCR-x proteins specifically bind to G proteins, they can be employed in any specific use where G protein specific affinity is required such as those uses where commercially available anti-G protein antibodies are employed.

Antibodies

5

10

15

20

25

30

8N50000 kW0 01-6473A2 1 8

Also comprehended by the present invention are antibodies (*e.g.*, monoclonal and polyclonal antibodies, single chain antibodies, chimeric antibodies, bifunctional/bispecific antibodies, humanized antibodies, human antibodies, and complementary determining region (CDR)-grafted antibodies, including compounds which include CDR sequences which specifically recognize a polypeptide of the invention) specific for nGPCR-x or fragments thereof. Preferred antibodies of the invention are human antibodies that are produced and identified according to methods described in WO93/11236, published June 20, 1993, which is incorporated herein by reference in its entirety. Antibody fragments, including Fab, Fab', F(ab')₂, and F_v, are also provided by the invention. The term "specific for," when used to describe antibodies of the invention, indicates that the variable regions of the antibodies of the

invention recognize and bind nGPCR-x polypeptides exclusively (*i.e.*, are able to distinguish nGPCR-x polypeptides from other known GPCR polypeptides by virtue of measurable differences in binding affinity, despite the possible existence of localized sequence identity, homology, or similarity between nGPCR-x and such polypeptides). It will be understood that specific antibodies may also interact with other proteins (for example, *S. aureus* protein A or other antibodies in ELISA techniques) through interactions with sequences outside the variable region of the antibodies, and, in particular, in the constant region of the molecule. Screening assays to determine binding specificity of an antibody of the invention are well known and routinely practiced in the art. For a comprehensive discussion of such assays, see Harlow *et al.* (Eds.), Antibodies A Laboratory Manual; Cold Spring Harbor Laboratory; Cold Spring Harbor, NY (1988), Chapter 6. Antibodies that recognize and bind fragments of the nGPCR-x polypeptides of the invention are also contemplated, provided that the antibodies are specific for nGPCR-x polypeptides. Antibodies of the invention can be produced using any method well known and routinely practiced in the art.

5

10

15

20

25

30

DOCID -WO 0136473A2 | 2

The invention provides an antibody that is specific for the nGPCR-x of the invention. Antibody specificity is described in greater detail below. However, it should be emphasized that antibodies that can be generated from polypeptides that have previously been described in the literature and that are capable of fortuitously cross-reacting with nGPCR-x (e.g., due to the fortuitous existence of a similar epitope in both polypeptides) are considered "cross-reactive" antibodies. Such cross-reactive antibodies are not antibodies that are "specific" for nGPCR-x. The determination of whether an antibody is specific for nGPCR-x or is cross-reactive with another known receptor is made using any of several assays, such as Western blotting assays, that are well known in the art. For identifying cells that express nGPCR-x and also for modulating nGPCR-x-ligand binding activity, antibodies that specifically bind to an extracellular epitope of the nGPCR-x are preferred.

In one preferred variation, the invention provides monoclonal antibodies. Hybridomas that produce such antibodies also are intended as aspects of the invention. In yet another variation, the invention provides a humanized antibody. Humanized antibodies are useful for *in vivo* therapeutic indications.

In another variation, the invention provides a cell-free composition comprising polyclonal antibodies, wherein at least one of the antibodies is an antibody of the invention specific for nGPCR-x. Antisera isolated from an animal is an exemplary

composition, as is a composition comprising an antibody fraction of an antisera that has been resuspended in water or in another diluent, excipient, or carrier.

In still another related embodiment, the invention provides an anti-idiotypic antibody specific for an antibody that is specific for nGPCR-x.

It is well known that antibodies contain relatively small antigen binding domains that can be isolated chemically or by recombinant techniques. Such domains are useful nGPCR-x binding molecules themselves, and also may be reintroduced into human antibodies, or fused to toxins or other polypeptides. Thus, in still another embodiment, the invention provides a polypeptide comprising a fragment of a nGPCR-x-specific antibody, wherein the fragment and the polypeptide bind to the nGPCR-x. By way of non-limiting example, the invention provides polypeptides that are single chain antibodies and CDR-grafted antibodies.

Non-human antibodies may be humanized by any of the methods known in the art. In one method, the non-human CDRs are inserted into a human antibody or consensus antibody framework sequence. Further changes can then be introduced into the antibody framework to modulate affinity or immunogenicity.

Antibodies of the invention are useful for, e.g., therapeutic purposes (by modulating activity of nGPCR-x), diagnostic purposes to detect or quantitate nGPCR-x, and purification of nGPCR-x. Kits comprising an antibody of the invention for any of the purposes described herein are also comprehended. In general, a kit of the invention also includes a control antigen for which the antibody is immunospecific.

Compositions

5

10

15

20

25

30

Mutations in the nGPCR-x gene that result in loss of normal function of the nGPCR-x gene product underlie nGPCR-x-related human disease states. The invention comprehends gene therapy to restore nGPCR-x activity to treat those disease states. Delivery of a functional nGPCR-x gene to appropriate cells is effected ex vivo, in situ, or in vivo by use of vectors, and more particularly viral vectors (e.g., adenovirus, adeno-associated virus, or a retrovirus), or ex vivo by use of physical DNA transfer methods (e.g., liposomes or chemical treatments). See, for example, Anderson, Nature, supplement to vol. 392, no. 6679, pp.25-20 (1998). For additional reviews of gene therapy technology see Friedmann, Science. 244: 1275-1281 (1989); Verma, Scientific American: 68-84 (1990); and Miller, Nature, 357: 455-460 (1992). Alternatively, it is contemplated that in other human disease states, preventing the expression of, or inhibiting the activity of, nGPCR-x will be useful in treating disease

states. It is contemplated that antisense therapy or gene therapy could be applied to negatively regulate the expression of nGPCR-x.

Another aspect of the present invention is directed to compositions, including pharmaceutical compositions, comprising any of the nucleic acid molecules or recombinant expression vectors described above and an acceptable carrier or diluent. Preferably, the carrier or diluent is pharmaceutically acceptable. Suitable carriers are described in the most recent edition of *Remington's Pharmaceutical Sciences*, A. Osol, a standard reference text in this field, which is incorporated herein by reference in its entirety. Preferred examples of such carriers or diluents include, but are not limited to, water, saline, Ringer's solution, dextrose solution, and 5% human serum albumin. Liposomes and nonaqueous vehicles such as fixed oils may also be used. The formulations are sterilized by commonly used techniques.

Also within the scope of the invention are compositions comprising polypeptides, polynucleotides, or antibodies of the invention that have been formulated with, *e.g.*, a pharmaceutically acceptable carrier.

10

15

20

25

30

BNSDOCID <WO 0136473A2 >

The invention also provides methods of using antibodies of the invention. For example, the invention provides a method for modulating ligand binding of a nGPCR-x comprising the step of contacting the nGPCR-x with an antibody specific for the nGPCR-x, under conditions wherein the antibody binds the receptor.

GPCRs that may be expressed in the brain, such as nGPCR-x, provide an indication that aberrant nGPCR-x signaling activity may correlate with one or more neurological or psychological disorders. The invention also provides a method for treating a neurological or psychiatric disorder comprising the step of administering to a mammal in need of such treatment an amount of an antibody-like polypeptide of the invention that is sufficient to modulate ligand binding to a nGPCR-x in neurons of the mammal. nGPCR-x may also be expressed in other tissues, including but not limited to, peripheral blood lymphocytes, pancreas, ovary, uterus, testis, salivary gland, thyroid gland, kidney, adrenal gland, liver, bone marrow, prostate, fetal liver, colon, muscle, and fetal brain, and may be found in many other tissues. Within the brain, nGPCR-x mRNA transcripts may be found in many tissues, including, but not limited to, frontal lobe, hypothalamus, pons, cerebellum, caudate nucleus, and medulla. Tissues and brain regions where specific nGPCRs of the present invention are expressed are identified in the Examples below.

Kits

5

10

15

20

25

30

BNSDD0 C kW0 - 0136473A2 F >

The present invention is also directed to kits, including pharmaceutical kits. The kits can comprise any of the nucleic acid molecules described above, any of the polypeptides described above, or any antibody which binds to a polypeptide of the invention as described above, as well as a negative control. The kit preferably comprises additional components, such as, for example, instructions, solid support, reagents helpful for quantification, and the like.

In another aspect, the invention features methods for detection of a polypeptide in a sample as a diagnostic tool for diseases or disorders, wherein the method comprises the steps of: (a) contacting the sample with a nucleic acid probe which hybridizes under hybridization assay conditions to a nucleic acid target region of a polypeptide having the sequence of even numbered sequences ranging from SEQ ID NO: 2 to SEQ ID NO: 94 and SEQ ID NO: 186, said probe comprising the nucleic acid sequence encoding the polypeptide, fragments thereof, and the complements of the sequences and fragments; and (b) detecting the presence or amount of the probe:target region hybrid as an indication of the disease.

In preferred embodiments of the invention, the disease is selected from the group consisting of thyroid disorders (e.g. thyreotoxicosis, myxoedema); renal failure; inflammatory conditions (e.g., Crohn's disease); diseases related to cell differentiation and homeostasis; rheumatoid arthritis; autoimmune disorders; movement disorders; CNS disorders (e.g., pain including migraine; stroke; psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, anxiety, generalized anxiety disorder, post-traumatic-stress disorder, depression, bipolar disorder, delirium, dementia, severe mental retardation; dyskinesias, such as Huntington's disease or Tourette's Syndrome; attention disorders including ADD and ADHD, and degenerative disorders such as Parkinson's, Alzheimer's; movement disorders, including ataxias, supranuclear palsy, etc.); infections, such as viral infections caused by HIV-1 or HIV-2; metabolic and cardiovascular diseases and disorders (e.g., type 2 diabetes, obesity, anorexia, hypotension, hypertension, thrombosis, myocardial infarction, cardiomyopathies, atherosclerosis, etc.); proliferative diseases and cancers (e.g., different cancers such as breast, colon, lung, etc., and hyperproliferative disorders such as psoriasis, prostate hyperplasia, etc.); hormonal disorders (e.g., male/female hormonal replacement, polycystic ovarian syndrome, alopecia, etc.); and sexual dysfunction, among others.

5

10

15

20

25

30

1-55/00/D kWC 0136473A2 (2

As described above and in Example 4 below, the genes encoding nGPCR-1 (nucleic acid sequence SEO ID NO: 1, SEO ID NO: 73, amino acid sequence SEO ID NO: 2, SEO ID NO:74), nGPCR-9 (nucleic acid sequence SEO ID NO:9, SEO ID NO:77, amino acid sequence SEQ ID NO:10, SEQ ID NO:78), nGPCR-11 (nucleic acid sequence SEQ ID NO:11, SEQ ID NO:79, amino acid sequence SEO ID NO:12, SEQ ID NO:80), nGPCR-16 (nucleic acid sequence SEQ ID NO: 21, SEQ ID NO:81, amino acid sequence SEO ID NO: 22, SEO ID NO:82), nGPCR-40 (nucleic acid sequence SEO ID NO:53, SEO ID NO:83, amino acid sequence SEO ID NO:54, SEO ID NO:84), nGPCR-54 (nucleic acid sequence SEO ID NO:59, SEO ID NO:85, amino acid sequence SEO ID NO:60, SEO ID NO: 86), nGPCR-56 (nucleic acid sequence SEO ID NO:63, SEQ ID NO:87, SEQ ID NO:89, amino acid sequence SEO ID NO:64, SEQ ID NO: 88, SEQ ID NO:90), nGPCR-58 (nucleic acid sequence SEQ ID NO:67, SEQ ID NO 91, SEQ ID NO:93, amino acid sequence SEQ ID NO:68, SEO ID NO: 92, SEO ID NO:94) and nGPCR-3 (nucleic acid sequence SEO ID NO:3, SEQ ID NO:185, amino acid sequence SEQ ID NO:4, SEQ ID NO: 186) have been detected in brain tissue indicating that these n-GPCR-x proteins are neuroreceptors. Kits may be designed to detect either expression of polynucleotides encoding these proteins or the proteins themselves in order to identify tissue as being neurological. For example, oligonucleotide hybridization kits can be provided which include a container having an oligonucleotide probe specific for the n-GPCR-xspecific DNA and optionally, containers with positive and negative controls and/or instructions. Similarly, PCR kits can be provided which include a container having primers specific for the n-GPCR-x-specific sequences, DNA and optionally. containers with size markers, positive and negative controls and/or instructions.

Hybridization conditions should be such that hybridization occurs only with the genes in the presence of other nucleic acid molecules. Under stringent hybridization conditions only highly complementary nucleic acid sequences hybridize. Preferably, such conditions prevent hybridization of nucleic acids having 1 or 2 mismatches out of 20 contiguous nucleotides. Such conditions are defined supra.

The diseases for which detection of genes in a sample could be diagnostic include diseases in which nucleic acid (DNA and/or RNA) is amplified in comparison to normal cells. By "amplification" is meant increased numbers of DNA or RNA in a cell compared with normal cells.

The diseases that could be diagnosed by detection of nucleic acid in a sample preferably include central nervous system and metabolic diseases. The test samples suitable for nucleic acid probing methods of the present invention include, for example, cells or nucleic acid extracts of cells, or biological fluids. The samples used in the above-described methods will vary based on the assay format, the detection method and the nature of the tissues, cells or extracts to be assayed. Methods for preparing nucleic acid extracts of cells are well known in the art and can be readily adapted in order to obtain a sample that is compatible with the method utilized.

Alternatively, immunoassay kits can be provided which have containers container having antibodies specific for the n-GPCR-x-protein and optionally, containers with positive and negative controls and/or instructions.

Kits may also be provided useful in the identification of GPCR binding partners such as natural ligands or modulators (agonists or antagonists). Substances useful for treatment of disorders or diseases preferably show positive results in one or more *in vitro* assays for an activity corresponding to treatment of the disease or disorder in question. Substances that modulate the activity of the polypeptides preferably include, but are not limited to, antisense oligonucleotides, agonists and antagonists, and inhibitors of protein kinases.

Methods of inducing immune response

Another aspect of the present invention is directed to methods of inducing an immune response in a mammal against a polypeptide of the invention by administering to the mammal an amount of the polypeptide sufficient to induce an immune response. The amount will be dependent on the animal species, size of the animal, and the like but can be determined by those skilled in the art.

Methods of identifying ligands

5

10

15

20

25

30

1-32473A2

BNSCOOT WY

The invention also provides assays to identify compounds that bind nGPCR-x. One such assay comprises the steps of: (a) contacting a composition comprising a nGPCR-x with a compound suspected of binding nGPCR-x; and (b) measuring binding between the compound and nGPCR-x. In one variation, the composition comprises a cell expressing nGPCR-x on its surface. In another variation, isolated nGPCR-x or cell membranes comprising nGPCR-x are employed. The binding may be measured directly, *e.g.*, by using a labeled compound, or may be measured indirectly by several techniques, including measuring intracellular signaling of

nGPCR-x induced by the compound (or measuring changes in the level of nGPCR-x signaling).

Specific binding molecules, including natural ligands and synthetic compounds, can be identified or developed using isolated or recombinant nGPCR-x products, nGPCR-x variants, or preferably, cells expressing such products. Binding partners are useful for purifying nGPCR-x products and detection or quantification of nGPCR-x products in fluid and tissue samples using known immunological procedures. Binding molecules are also manifestly useful in modulating (*i.e.*, blocking, inhibiting or stimulating) biological activities of nGPCR-x, especially those activities involved in signal transduction.

The DNA and amino acid sequence information provided by the present invention also makes possible identification of binding partner compounds with which a nGPCR-x polypeptide or polynucleotide will interact. Methods to identify binding partner compounds include solution assays, *in vitro* assays wherein nGPCR-x polypeptides are immobilized, and cell-based assays. Identification of binding partner compounds of nGPCR-x polypeptides provides candidates for therapeutic or prophylactic intervention in pathologies associated with nGPCR-x normal and aberrant biological activity.

The invention includes several assay systems for identifying nGPCR-x binding partners. In solution assays, methods of the invention comprise the steps of (a) contacting a nGPCR-x polypeptide with one or more candidate binding partner compounds and (b) identifying the compounds that bind to the nGPCR-x polypeptide. Identification of the compounds that bind the nGPCR-x polypeptide can be achieved by isolating the nGPCR-x polypeptide/binding partner complex, and separating the binding partner compound from the nGPCR-x polypeptide. An additional step of characterizing the physical, biological, and/or biochemical properties of the binding partner compound is also comprehended in another embodiment of the invention. In one aspect, the nGPCR-x polypeptide/binding partner complex is isolated using an antibody immunospecific for either the nGPCR-x polypeptide or the candidate binding partner compound.

In still other embodiments, either the nGPCR-x polypeptide or the candidate binding partner compound comprises a label or tag that facilitates its isolation, and methods of the invention to identify binding partner compounds include a step of isolating the nGPCR-x polypeptide/binding partner complex through interaction with

5

10

15

20

25

the label or tag. An exemplary tag of this type is a poly-histidine sequence, generally around six histidine residues, that permits isolation of a compound so labeled using nickel chelation. Other labels and tags, such as the FLAG® tag (Eastman Kodak, Rochester, NY), well known and routinely used in the art, are embraced by the invention.

5

10

15

20

25

30

NSD0015 9WC - 136475A2 1 8

In one variation of an *in vitro* assay, the invention provides a method comprising the steps of (a) contacting an immobilized nGPCR-x polypeptide with a candidate binding partner compound and (b) detecting binding of the candidate compound to the nGPCR-x polypeptide. In an alternative embodiment, the candidate binding partner compound is immobilized and binding of nGPCR-x is detected. Immobilization is accomplished using any of the methods well known in the art, including covalent bonding to a support, a bead, or a chromatographic resin, as well as non-covalent, high affinity interactions such as antibody binding, or use of streptavidin/biotin binding wherein the immobilized compound includes a biotin moiety. Detection of binding can be accomplished (i) using a radioactive label on the compound that is not immobilized, (ii) using of a fluorescent label on the non-immobilized compound, (iii) using an antibody immunospecific for the non-immobilized compound, (iv) using a label on the non-immobilized compound that excites a fluorescent support to which the immobilized compound is attached, as well as other techniques well known and routinely practiced in the art.

The invention also provides cell-based assays to identify binding partner compounds of a nGPCR-x polypeptide. In one embodiment, the invention provides a method comprising the steps of contacting a nGPCR-x polypeptide expressed on the surface of a cell with a candidate binding partner compound and detecting binding of the candidate binding partner compound to the nGPCR-x polypeptide. In a preferred embodiment, the detection comprises detecting a calcium flux or other physiological event in the cell caused by the binding of the molecule.

Another aspect of the present invention is directed to methods of identifying compounds that bind to either nGPCR-x or nucleic acid molecules encoding nGPCR-x, comprising contacting nGPCR-x, or a nucleic acid molecule encoding the same, with a compound, and determining whether the compound binds nGPCR-x or a nucleic acid molecule encoding the same. Binding can be determined by binding assays which are well known to the skilled artisan, including, but not limited to, gelshift assays, Western blots, radiolabeled competition assay, phage-based expression

cloning, co-fractionation by chromatography, co-precipitation, cross linking. interaction trap/two-hybrid analysis, southwestern analysis, ELISA, and the like, which are described in, for example, Current Protocols in Molecular Biology, 1999, John Wiley & Sons, NY, which is incorporated herein by reference in its entirety. The compounds to be screened include (which may include compounds which are suspected to bind nGPCR-x, or a nucleic acid molecule encoding the same), but are not limited to, extracellular, intracellular, biologic or chemical origin. The methods of the invention also embrace ligands, especially neuropeptides, that are attached to a label, such as a radiolabel (e.g., 125I, 35S, 32P, 33P, 3H), a fluorescence label, a chemiluminescent label, an enzymic label and an immunogenic label. Modulators falling within the scope of the invention include, but are not limited to, non-peptide molecules such as non-peptide mimetics, non-peptide allosteric effectors, and peptides. The nGPCR-x polypeptide or polynucleotide employed in such a test may either be free in solution, attached to a solid support, borne on a cell surface or located intracellularly or associated with a portion of a cell. One skilled in the art can, for example, measure the formation of complexes between nGPCR-x and the compound being tested. Alternatively, one skilled in the art can examine the diminution in complex formation between nGPCR-x and its substrate caused by the compound being tested.

10

15

20

25

30

In another embodiment of the invention, high throughput screening for compounds having suitable binding affinity to nGPCR-x is employed. Briefly, large numbers of different small peptide test compounds are synthesized on a solid substrate. The peptide test compounds are contacted with nGPCR-x and washed. Bound nGPCR-x is then detected by methods well known in the art. Purified polypeptides of the invention can also be coated directly onto plates for use in the aforementioned drug screening techniques. In addition, non-neutralizing antibodies can be used to capture the protein and immobilize it on the solid support.

Generally, an expressed nGPCR-x can be used for HTS binding assays in conjunction with its defined ligand, in this case the corresponding neuropeptide that activates it. The identified peptide is labeled with a suitable radioisotope, including, but not limited to, ¹²⁵I, ³H, ³⁵S or ³²P, by methods that are well known to those skilled in the art. Alternatively, the peptides may be labeled by well-known methods with a suitable fluorescent derivative (Baindur *et al.*, *Drug Dev. Res.*, **1994**, *33*, 373-398; Rogers, *Drug Discovery Today*, **1997**, *2*, 156-160). Radioactive ligand specifically

bound to the receptor in membrane preparations made from the cell line expressing the recombinant protein can be detected in HTS assays in one of several standard ways, including filtration of the receptor-ligand complex to separate bound ligand from unbound ligand (Williams, Med. Res. Rev., 1991, 11, 147-184; Sweetnam et al., J. Natural Products, 1993, 56, 441-455). Alternative methods include a scintillation proximity assay (SPA) or a FlashPlate format in which such separation is unnecessary (Nakayama, Cur. Opinion Drug Disc. Dev., 1998, 1, 85-91 Bossé et al., J. Biomolecular Screening, 1998, 3, 285-292.). Binding of fluorescent ligands can be detected in various ways, including fluorescence energy transfer (FRET), direct spectrophotofluorometric analysis of bound ligand, or fluorescence polarization (Rogers, Drug Discovery Today, 1997, 2, 156-160; Hill, Cur. Opinion Drug Disc. Dev., 1998, 1, 92-97).

Other assays may be used to identify specific ligands of a nGPCR-x receptor, including assays that identify ligands of the target protein through measuring direct binding of test ligands to the target protein, as well as assays that identify ligands of target proteins through affinity ultrafiltration with ion spray mass spectroscopy/HPLC methods or other physical and analytical methods. Alternatively, such binding interactions are evaluated indirectly using the yeast two-hybrid system described in Fields et al., Nature, 340:245-246 (1989), and Fields et al., Trends in Genetics, 10:286-292 (1994), both of which are incorporated herein by reference. The twohybrid system is a genetic assay for detecting interactions between two proteins or polypeptides. It can be used to identify proteins that bind to a known protein of interest, or to delineate domains or residues critical for an interaction. Variations on this methodology have been developed to clone genes that encode DNA binding proteins, to identify peptides that bind to a protein, and to screen for drugs. The twohybrid system exploits the ability of a pair of interacting proteins to bring a transcription activation domain into close proximity with a DNA binding domain that binds to an upstream activation sequence (UAS) of a reporter gene, and is generally performed in yeast. The assay requires the construction of two hybrid genes encoding (1) a DNA-binding domain that is fused to a first protein and (2) an activation domain fused to a second protein. The DNA-binding domain targets the first hybrid protein to the UAS of the reporter gene; however, because most proteins lack an activation domain, this DNA-binding hybrid protein does not activate transcription of the reporter gene. The second hybrid protein, which contains the activation domain,

5

10

15

20

25

cannot by itself activate expression of the reporter gene because it does not bind the UAS. However, when both hybrid proteins are present, the noncovalent interaction of the first and second proteins tethers the activation domain to the UAS, activating transcription of the reporter gene. For example, when the first protein is a GPCR gene product, or fragment thereof, that is known to interact with another protein or nucleic acid, this assay can be used to detect agents that interfere with the binding interaction. Expression of the reporter gene is monitored as different test agents are added to the system. The presence of an inhibitory agent results in lack of a reporter signal.

10

15

20

25

30

BNSDOCID < WO 0136473A2 1 >

The function of nGPCR-x gene products is unclear and no ligands have yet been found which bind the gene product. The yeast two-hybrid assay can also be used to identify proteins that bind to the gene product. In an assay to identify proteins that bind to a nGPCR-x receptor, or fragment thereof, a fusion polynucleotide encoding both a nGPCR-x receptor (or fragment) and a UAS binding domain (*i.e.*, a first protein) may be used. In addition, a large number of hybrid genes each encoding a different second protein fused to an activation domain are produced and screened in the assay. Typically, the second protein is encoded by one or more members of a total cDNA or genomic DNA fusion library, with each second protein-coding region being fused to the activation domain. This system is applicable to a wide variety of proteins, and it is not even necessary to know the identity or function of the second binding protein. The system is highly sensitive and can detect interactions not revealed by other methods; even transient interactions may trigger transcription to produce a stable mRNA that can be repeatedly translated to yield the reporter protein.

Other assays may be used to search for agents that bind to the target protein. One such screening method to identify direct binding of test ligands to a target protein is described in U.S. Patent No. 5,585,277, incorporated herein by reference. This method relies on the principle that proteins generally exist as a mixture of folded and unfolded states, and continually alternate between the two states. When a test ligand binds to the folded form of a target protein (i.e., when the test ligand is a ligand of the target protein), the target protein molecule bound by the ligand remains in its folded state. Thus, the folded target protein is present to a greater extent in the presence of a test ligand which binds the target protein, than in the absence of a ligand. Binding of the ligand to the target protein can be determined by any method that distinguishes between the folded and unfolded states of the target protein. The function of the

target protein need not be known in order for this assay to be performed. Virtually any agent can be assessed by this method as a test ligand, including, but not limited to, metals, polypeptides, proteins, lipids, polysaccharides, polynucleotides and small organic molecules.

Another method for identifying ligands of a target protein is described in Wieboldt *et al.*, Anal. Chem., 69:1683-1691 (1997), incorporated herein by reference. This technique screens combinatorial libraries of 20-30 agents at a time in solution phase for binding to the target protein. Agents that bind to the target protein are separated from other library components by simple membrane washing. The specifically selected molecules that are retained on the filter are subsequently liberated from the target protein and analyzed by HPLC and pneumatically assisted electrospray (ion spray) ionization mass spectroscopy. This procedure selects library components with the greatest affinity for the target protein, and is particularly useful for small molecule libraries.

Other embodiments of the invention comprise using competitive screening assays in which neutralizing antibodies capable of binding a polypeptide of the invention specifically compete with a test compound for binding to the polypeptide. In this manner, the antibodies can be used to detect the presence of any peptide that shares one or more antigenic determinants with nGPCR-x. Radiolabeled competitive binding studies are described in A.H. Lin et al. Antimicrobial Agents and Chemotherapy, 1997, vol. 41, no. 10. pp. 2127-2131, the disclosure of which is incorporated herein by reference in its entirety.

As described above and in Example 4 below, the genes encoding nGPCR-1 (nucleic acid sequence SEQ ID NO: 1, SEQ ID NO: 73, amino acid sequence SEQ ID NO: 2, SEQ ID NO:74), nGPCR-9 (nucleic acid sequence SEQ ID NO:9, SEQ ID NO:77, amino acid sequence SEQ ID NO:10, SEQ ID NO:78), nGPCR-11 (nucleic acid sequence SEQ ID NO:11, SEQ ID NO:79, amino acid sequence SEQ ID NO:12, SEQ ID NO:80), nGPCR-16 (nucleic acid sequence SEQ ID NO: 21, SEQ ID NO:81, amino acid sequence SEQ ID NO: 22, SEQ ID NO:82), nGPCR-40 (nucleic acid sequence SEQ ID NO:53, SEQ ID NO:83, amino acid sequence SEQ ID NO:54, SEQ ID NO:84), nGPCR-54 (nucleic acid sequence SEQ ID NO:59, SEQ ID NO:85, amino acid sequence SEQ ID NO:60, SEQ ID NO:86), nGPCR-56 (nucleic acid sequence SEQ ID NO:63, SEQ ID NO:87, SEQ ID NO:89, amino acid sequence SEQ ID NO:64, SEQ ID NO:88, SEQ ID NO:90), nGPCR-58 (nucleic acid sequence SEQ ID NO:85, SEQ ID NO:90), nGPCR-58 (nucleic acid sequence SEQ ID NO:84, SEQ ID NO:88, SEQ ID NO:90), nGPCR-58 (nucleic acid sequence SEQ ID NO:64, SEQ ID NO:88, SEQ ID NO:90), nGPCR-58 (nucleic acid sequence SEQ ID NO:80, SEQ ID NO:90), nGPCR-58 (nucleic acid sequence SEQ ID NO:80, SEQ ID NO:90), nGPCR-58 (nucleic acid sequence SEQ ID NO:64, SEQ ID NO:88, SEQ ID NO:90), nGPCR-58 (nucleic acid sequence SEQ

5

10

15

20

25

ID NO:67, SEQ ID NO:91, SEQ ID NO:93, amino acid sequence SEQ ID NO:68, SEQ ID NO: 92, SEQ ID NO:94), and nGPCR-3 (nucleic acid sequence SEQ ID NO:3, SEQ ID NO:185, amino acid sequence SEQ ID NO:4, SEQ ID NO: 186) have been detected in brain tissue indicating that these n-GPCR-x proteins are neuroreceptors. Accordingly, natural binding partners of these molecules include neurotransmitters.

Identification of modulating agents

5

10

15

20

25

30

The invention also provides methods for identifying a modulator of binding between a nGPCR-x and a nGPCR-x binding partner, comprising the steps of: (a) contacting a nGPCR-x binding partner and a composition comprising a nGPCR-x in the presence and in the absence of a putative modulator compound; (b) detecting binding between the binding partner and the nGPCR-x; and (c) identifying a putative modulator compound or a modulator compound in view of decreased or increased binding between the binding partner and the nGPCR-x in the presence of the putative modulator, as compared to binding in the absence of the putative modulator.

nGPCR-x binding partners that stimulate nGPCR-x activity are useful as agonists in disease states or conditions characterized by insufficient nGPCR-x signaling (e.g., as a result of insufficient activity of a nGPCR-x ligand). nGPCR-x binding partners that block ligand-mediated nGPCR-x signaling are useful as nGPCR-x antagonists to treat disease states or conditions characterized by excessive nGPCR-x signaling. In addition nGPCR-x modulators in general, as well as nGPCR-x polynucleotides and polypeptides, are useful in diagnostic assays for such diseases or conditions.

In another aspect, the invention provides methods for treating a disease or abnormal condition by administering to a patient in need of such treatment a substance that modulates the activity or expression of a polypeptide having the sequence of even numbered sequences ranging from SEQ ID NO: 2 to SEQ ID NO: 94 and SEQ ID NO: 186.

Agents that modulate (*i.e.*, increase, decrease, or block) nGPCR-x activity or expression may be identified by incubating a putative modulator with a cell containing a nGPCR-x polypeptide or polynucleotide and determining the effect of the putative modulator on nGPCR-x activity or expression. The selectivity of a compound that modulates the activity of nGPCR-x can be evaluated by comparing its effects on nGPCR-x to its effect on other GPCR compounds. Selective modulators

may include, for example, antibodies and other proteins, peptides, or organic molecules that specifically bind to a nGPCR-x polypeptide or a nGPCR-x-encoding nucleic acid. Modulators of nGPCR-x activity will be therapeutically useful in treatment of diseases and physiological conditions in which normal or aberrant nGPCR-x activity is involved. nGPCR-x polynucleotides, polypeptides, and modulators may be used in the treatment of such diseases and conditions as infections, such as viral infections caused by HIV-1 or HIV-2; pain; cancers; Parkinson's disease; hypotension; hypertension; and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Tourette's Syndrome, among others. nGPCR-x polynucleotides and polypeptides, as well as nGPCR-x modulators, may also be used in diagnostic assays for such diseases or conditions.

Methods of the invention to identify modulators include variations on any of the methods described above to identify binding partner compounds, the variations including techniques wherein a binding partner compound has been identified and the binding assay is carried out in the presence and absence of a candidate modulator. A modulator is identified in those instances where binding between the nGPCR-x polypeptide and the binding partner compound changes in the presence of the candidate modulator compared to binding in the absence of the candidate modulator compound. A modulator that increases binding between the nGPCR-x polypeptide and the binding partner compound is described as an enhancer or activator, and a modulator that decreases binding between the nGPCR-x polypeptide and the binding partner compound is described as an inhibitor.

The invention also comprehends high-throughput screening (HTS) assays to identify compounds that interact with or inhibit biological activity (*i.e.*, affect enzymatic activity, binding activity, *etc.*) of a nGPCR-x polypeptide. HTS assays permit screening of large numbers of compounds in an efficient manner. Cell-based HTS systems are contemplated to investigate nGPCR-x receptor-ligand interaction. HTS assays are designed to identify "hits" or "lead compounds" having the desired property, from which modifications can be designed to improve the desired property. Chemical modification of the "hit" or "lead compound" is often based on an identifiable structure/activity relationship between the "hit" and the nGPCR-x polypeptide.

5

10

15

20

25

Another aspect of the present invention is directed to methods of identifying compounds which modulate (i.e., increase or decrease) activity of nGPCR-x comprising contacting nGPCR-x with a compound, and determining whether the compound modifies activity of nGPCR-x. The activity in the presence of the test compared is measured to the activity in the absence of the test compound. Where the activity of the sample containing the test compound is higher than the activity in the sample lacking the test compound, the compound will have increased activity. Similarly, where the activity of the sample containing the test compound is lower than the activity in the sample lacking the test compound, the compound will have inhibited activity.

5

10

15

20

25

30

SDOCID #WO 0136473A2 La

The present invention is particularly useful for screening compounds by using nGPCR-x in any of a variety of drug screening techniques. The compounds to be screened include (which may include compounds which are suspected to modulate nGPCR-x activity), but are not limited to, extracellular, intracellular, biologic or chemical origin. The nGPCR-x polypeptide employed in such a test may be in any form, preferably, free in solution, attached to a solid support, borne on a cell surface or located intracellularly. One skilled in the art can, for example, measure the formation of complexes between nGPCR-x and the compound being tested.

Alternatively, one skilled in the art can examine the diminution in complex formation between nGPCR-x and its substrate caused by the compound being tested.

The activity of nGPCR-x polypeptides of the invention can be determined by, for example, examining the ability to bind or be activated by chemically synthesized peptide ligands. Alternatively, the activity of nGPCR-x polypeptides can be assayed by examining their ability to bind calcium ions, hormones, chemokines, neuropeptides, neurotransmitters, nucleotides, lipids, odorants, and photons. Alternatively, the activity of the nGPCR-x polypeptides can be determined by examining the activity of effector molecules including, but not limited to, adenylate cyclase, phospholipases and ion channels. Thus, modulators of nGPCR-x polypeptide activity may alter a GPCR receptor function, such as a binding property of a receptor or an activity such as G protein-mediated signal transduction or membrane localization. In various embodiments of the method, the assay may take the form of an ion flux assay, a yeast growth assay, a non-hydrolyzable GTP assay such as a [35S]-GTP S assay, a cAMP assay, an inositol triphosphate assay, a diacylglycerol assay, an Aequorin assay, a Luciferase assay, a FLIPR assay for intracellular Ca²⁺

concentration, a mitogenesis assay, a MAP Kinase activity assay, an arachidonic acid release assay (e.g., using [³H]-arachidonic acid), and an assay for extracellular acidification rates, as well as other binding or function-based assays of nGPCR-x activity that are generally known in the art. In several of these embodiments, the invention comprehends the inclusion of any of the G proteins known in the art, such as G 16, G 15, or chimeric Gqd5, Gq55, Gq05, Gq25, and the like. nGPCR-x activity can be determined by methodologies that are used to assay for FaRP activity, which is well known to those skilled in the art. Biological activities of nGPCR-x receptors according to the invention include, but are not limited to, the binding of a natural or an unnatural ligand, as well as any one of the functional activities of GPCRs known in the art. Non-limiting examples of GPCR activities include transmembrane signaling of various forms, which may involve G protein association and/or the exertion of an influence over G protein binding of various guanidylate nucleotides; another exemplary activity of GPCRs is the binding of accessory proteins or polypeptides that differ from known G proteins.

The modulators of the invention exhibit a variety of chemical structures, which can be generally grouped into non-peptide mimetics of natural GPCR receptor ligands, peptide and non-peptide allosteric effectors of GPCR receptors, and peptides that may function as activators or inhibitors (competitive, uncompetitive and non-competitive) (e.g., antibody products) of GPCR receptors. The invention does not restrict the sources for suitable modulators, which may be obtained from natural sources such as plant, animal or mineral extracts, or non-natural sources such as small molecule libraries, including the products of combinatorial chemical approaches to library construction, and peptide libraries. Examples of peptide modulators of GPCR receptors exhibit the following primary structures: GLGPRPLRFamide, GNSFLRFamide, GGPQGPLRFamide, GPSGPLRFamide, PDVDHVFLRFamide, and pyro-EDVDHVFLRFamide.

Other assays can be used to examine enzymatic activity including, but not limited to, photometric, radiometric, HPLC, electrochemical, and the like, which are described in, for example, *Enzyme Assays: A Practical Approach*, eds. R. Eisenthal and M. J. Danson, 1992, Oxford University Press, which is incorporated herein by reference in its entirety.

The use of cDNAs encoding GPCRs in drug discovery programs is well-known; assays capable of testing thousands of unknown compounds per day in high-

30

5

10

15

throughput screens (HTSs) are thoroughly documented. The literature is replete with examples of the use of radiolabelled ligands in HTS binding assays for drug discovery (see Williams, Medicinal Research Reviews, 1991, 11, 147-184.; Sweetnam, et al., J. Natural Products, 1993, 56, 441-455 for review). Recombinant receptors are preferred for binding assay HTS because they allow for better specificity (higher relative purity), provide the ability to generate large amounts of receptor material, and can be used in a broad variety of formats (see Hodgson, Bio/Technology, 1992, 10, 973-980; each of which is incorporated herein by reference in its entirety).

A variety of heterologous systems is available for functional expression of recombinant receptors that are well known to those skilled in the art. Such systems include bacteria (Strosberg, et al., Trends in Pharmacological Sciences, 1992, 13, 95-98), yeast (Pausch, Trends in Biotechnology, 1997, 15, 487-494), several kinds of insect cells (Vanden Broeck, Int. Rev. Cytology, 1996, 164, 189-268), amphibian cells (Jayawickreme et al., Current Opinion in Biotechnology, 1997, 8, 629-634) and several mammalian cell lines (CHO, HEK293, COS, etc.; see Gerhardt, et al., Eur. J. Pharmacology, 1997, 334, 1-23). These examples do not preclude the use of other possible cell expression systems, including cell lines obtained from nematodes (PCT application WO 98/37177).

In preferred embodiments of the invention, methods of screening for compounds that modulate nGPCR-x activity comprise contacting test compounds with nGPCR-x and assaying for the presence of a complex between the compound and nGPCR-x. In such assays, the ligand is typically labeled. After suitable incubation, free ligand is separated from that present in bound form, and the amount of free or uncomplexed label is a measure of the ability of the particular compound to bind to nGPCR-x.

It is well known that activation of heterologous receptors expressed in recombinant systems results in a variety of biological responses, which are mediated by G proteins expressed in the host cells. Occupation of a GPCR by an agonist results in exchange of bound GDP for GTP at a binding site on the G_{α} subunit; one can use a radioactive, non-hydrolyzable derivative of GTP, GTP χ^{35} S], to measure binding of an agonist to the receptor (Sim *et al.*, Neuroreport, 1996, 7, 729-733). One can also use this binding to measure the ability of antagonists to bind to the receptor by decreasing binding of GTP χ^{35} S] in the presence of a known agonist. One could

5

10

15

20

25

therefore construct a HTS based on GTP χ [35S] binding, though this is not the preferred method.

The G proteins required for functional expression of heterologous GPCRs can be native constituents of the host cell or can be introduced through well-known recombinant technology. The G proteins can be intact or chimeric. Often, a nearly universally competent G protein (e.g., $G_{\alpha 16}$) is used to couple any given receptor to a detectable response pathway. G protein activation results in the stimulation or inhibition of other native proteins, events that can be linked to a measurable response.

Examples of such biological responses include, but are not limited to, the following: the ability to survive in the absence of a limiting nutrient in specifically engineered yeast cells (Pausch, *Trends in Biotechnology*, 1997, 15, 487-494); changes in intracellular Ca²⁺ concentration as measured by fluorescent dyes (Murphy, et al., Cur. Opinion Drug Disc. Dev., 1998, 1, 192-199). Fluorescence changes can also be used to monitor ligand-induced changes in membrane potential or intracellular pH; an automated system suitable for HTS has been described for these purposes (Schroeder, et al., J. Biomolecular Screening, 1996, 1, 75-80). Melanophores prepared from Xenopus laevis show a ligand-dependent change in pigment organization in response to heterologous GPCR activation; this response is adaptable to HTS formats (Jayawickreme et al., Cur. Opinion Biotechnology, 1997, 8, 629-634). Assays are also available for the measurement of common second messengers, including cAMP, phosphoinositides and arachidonic acid, but these are not generally preferred for HTS.

Preferred methods of HTS employing these receptors include permanently transfected CHO cells, in which agonists and antagonists can be identified by the ability to specifically alter the binding of $GTP\gamma[^{35}S]$ in membranes prepared from these cells. In another embodiment of the invention, permanently transfected CHO cells could be used for the preparation of membranes which contain significant amounts of the recombinant receptor proteins; these membrane preparations would then be used in receptor binding assays, employing the radiolabelled ligand specific for the particular receptor. Alternatively, a functional assay, such as fluorescent monitoring of ligand-induced changes in internal Ca^{2+} concentration or membrane potential in permanently transfected CHO cells containing each of these receptors individually or in combination would be preferred for HTS. Equally preferred would be an alternative type of mammalian cell, such as HEK293 or COS cells, in similar

5

10

15

20

25

formats. More preferred would be permanently transfected insect cell lines, such as *Drosophila* S2 cells. Even more preferred would be recombinant yeast cells expressing the *Drosophila melanogaster* receptors in HTS formats well known to those skilled in the art (e.g., Pausch, *Trends in Biotechnology*, 1997, 15, 487-494).

The invention contemplates a multitude of assays to screen and identify inhibitors of ligand binding to nGPCR-x receptors. In one example, the nGPCR-x receptor is immobilized and interaction with a binding partner is assessed in the presence and absence of a candidate modulator such as an inhibitor compound. In another example, interaction between the nGPCR-x receptor and its binding partner is assessed in a solution assay, both in the presence and absence of a candidate inhibitor compound. In either assay, an inhibitor is identified as a compound that decreases binding between the nGPCR-x receptor and its binding partner. Another contemplated assay involves a variation of the dihybrid assay wherein an inhibitor of protein/protein interactions is identified by detection of a positive signal in a transformed or transfected host cell, as described in PCT publication number WO 95/20652, published August 3, 1995.

Candidate modulators contemplated by the invention include compounds selected from libraries of either potential activators or potential inhibitors. There are a number of different libraries used for the identification of small molecule modulators, including: (1) chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of random peptides, oligonucleotides or organic molecules. Chemical libraries consist of random chemical structures, some of which are analogs of known compounds or analogs of compounds that have been identified as "hits" or "leads" in other drug discovery screens, some of which are derived from natural products, and some of which arise from non-directed synthetic organic chemistry. Natural product libraries are collections of microorganisms, animals, plants, or marine organisms which are used to create mixtures for screening by: (1) fermentation and extraction of broths from soil, plant or marine microorganisms or (2) extraction of plants or marine organisms. Natural product libraries include polyketides, nonribosomal peptides, and variants (non-naturally occurring) thereof. For a review, see Science 282:63-68 (1998). Combinatorial libraries are composed of large numbers of peptides, oligonucleotides, or organic compounds as a mixture. These libraries are relatively easy to prepare by traditional automated synthesis methods, PCR, cloning, or proprietary synthetic methods. Of particular interest are non-peptide combinatorial

5

10

15

20

25

libraries. Still other libraries of interest include peptide, protein, peptidomimetic, multiparallel synthetic collection, recombinatorial, and polypeptide libraries. For a review of combinatorial chemistry and libraries created therefrom, see Myers, Curr. Opin. Biotechnol. 8:701-707 (1997). Identification of modulators through use of the various libraries described herein permits modification of the candidate "hit" (or "lead") to optimize the capacity of the "hit" to modulate activity.

Still other candidate inhibitors contemplated by the invention can be designed and include soluble forms of binding partners, as well as such binding partners as chimeric, or fusion, proteins. A "binding partner" as used herein broadly encompasses non-peptide modulators, as well as such peptide modulators as neuropeptides other than natural ligands, antibodies, antibody fragments, and modified compounds comprising antibody domains that are immunospecific for the expression product of the identified nGPCR-x gene.

The polypeptides of the invention are employed as a research tool for identification, characterization and purification of interacting, regulatory proteins. Appropriate labels are incorporated into the polypeptides of the invention by various methods known in the art and the polypeptides are used to capture interacting molecules. For example, molecules are incubated with the labeled polypeptides, washed to remove unbound polypeptides, and the polypeptide complex is quantified. Data obtained using different concentrations of polypeptide are used to calculate values for the number, affinity, and association of polypeptide with the protein complex.

Labeled polypeptides are also useful as reagents for the purification of molecules with which the polypeptide interacts including, but not limited to, inhibitors. In one embodiment of affinity purification, a polypeptide is covalently coupled to a chromatography column. Cells and their membranes are extracted, and various cellular subcomponents are passed over the column. Molecules bind to the column by virtue of their affinity to the polypeptide. The polypeptide-complex is recovered from the column, dissociated and the recovered molecule is subjected to protein sequencing. This amino acid sequence is then used to identify the captured molecule or to design degenerate oligonucleotides for cloning the corresponding gene from an appropriate cDNA library.

Alternatively, compounds may be identified which exhibit similar properties to the ligand for the nGPCR-x of the invention, but which are smaller and exhibit a

5

10

15

20

25

longer half time than the endogenous ligand in a human or animal body. When an organic compound is designed, a molecule according to the invention is used as a "lead" compound. The design of mimetics to known pharmaceutically active compounds is a well-known approach in the development of pharmaceuticals based on such "lead" compounds. Mimetic design, synthesis and testing are generally used to avoid randomly screening a large number of molecules for a target property. Furthermore, structural data deriving from the analysis of the deduced amino acid sequences encoded by the DNAs of the present invention are useful to design new drugs, more specific and therefore with a higher pharmacological potency.

5

10

15

20

25

30

DOCID 4WO 0136473A2 | 5

Comparison of the protein sequence of the present invention with the sequences present in all the available databases showed a significant homology with the transmembrane portion of G protein coupled receptors. Accordingly, computer modeling can be used to develop a putative tertiary structure of the proteins of the invention based on the available information of the transmembrane domain of other proteins. Thus, novel ligands based on the predicted structure of nGPCR-x can be designed.

In a particular embodiment, the novel molecules identified by the screening methods according to the invention are low molecular weight organic molecules, in which case a composition or pharmaceutical composition can be prepared thereof for oral intake, such as in tablets. The compositions, or pharmaceutical compositions, comprising the nucleic acid molecules, vectors, polypeptides, antibodies and compounds identified by the screening methods described herein, can be prepared for any route of administration including, but not limited to, oral, intravenous, cutaneous, subcutaneous, nasal, intramuscular or intraperitoneal. The nature of the carrier or other ingredients will depend on the specific route of administration and particular embodiment of the invention to be administered. Examples of techniques and protocols that are useful in this context are, *inter alia*, found in Remington's Pharmaceutical Sciences, 16th edition, Osol, A (ed.), 1980, which is incorporated herein by reference in its entirety.

The dosage of these low molecular weight compounds will depend on the disease state or condition to be treated and other clinical factors such as weight and condition of the human or animal and the route of administration of the compound. For treating human or animals, between approximately 0.5 mg/kg of body weight to 500 mg/kg of body weight of the compound can be administered. Therapy is

typically administered at lower dosages and is continued until the desired therapeutic outcome is observed.

The present compounds and methods, including nucleic acid molecules, polypeptides, antibodies, compounds identified by the screening methods described herein, have a variety of pharmaceutical applications and may be used, for example, to treat or prevent unregulated cellular growth, such as cancer cell and tumor growth. In a particular embodiment, the present molecules are used in gene therapy. For a review of gene therapy procedures, see *e.g.* Anderson, *Science*, 1992, 256, 808-813, which is incorporated herein by reference in its entirety.

5

10

15

20

25

30

36473A2 L >

The present invention also encompasses a method of agonizing (stimulating) or antagonizing a nGPCR-x natural binding partner associated activity in a mammal comprising administering to said mammal an agonist or antagonist to one of the above disclosed polypeptides in an amount sufficient to effect said agonism or antagonism. One embodiment of the present invention, then, is a method of treating diseases in a mammal with an agonist or antagonist of the protein of the present invention comprises administering the agonist or antagonist to a mammal in an amount sufficient to agonize or antagonize nGPCR-x-associated functions.

In an effort to discover novel treatments for diseases, biomedical researchers and chemists have designed, synthesized, and tested molecules that inhibit the function of protein polypeptides. Some small organic molecules form a class of compounds that modulate the function of protein polypeptides. Examples of molecules that have been reported to inhibit the function of protein kinases include, but are not limited to, bis monocyclic, bicyclic or heterocyclic aryl compounds (PCT WO 92/20642, published November 26, 1992 by Maguire et al.), vinylene-azaindole derivatives (PCT WO 94/14808, published July 7, 1994 by Ballinari et al.), 1cyclopropyl-4-pyridyl-quinolones (U.S. Patent No. 5,330,992), styryl compounds (U.S. Patent No. 5,217,999), styryl-substituted pyridyl compounds (U.S. Patent No. 5,302,606), certain quinazoline derivatives (EP Application No. 0 566 266 A1), seleoindoles and selenides (PCT WO 94/03427, published February 17, 1994 by Denny et al.), tricyclic polyhydroxylic compounds (PCT WO 92/21660, published December 10, 1992 by Dow), and benzylphosphonic acid compounds (PCT WO 91/15495, published October 17, 1991 by Dow et al), all of which are incorporated by reference herein, including any drawings.

Exemplary diseases and conditions amenable to treatment based on the present invention include, but are not limited to, thyroid disorders (e.g. thyreotoxicosis, myxoedema); renal failure; inflammatory conditions (e.g., Chron's disease); diseases related to cell differentiation and homeostasis; rheumatoid arthritis; autoimmune disorders; movement disorders; CNS disorders (e.g., pain including migraine; stroke; psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, anxiety, generalized anxiety disorder, post-traumatic-stress disorder, depression, bipolar disorder, delirium, dementia, severe mental retardation; dyskinesias, such as Huntington's disease or Tourette's Syndrome; attention disorders including ADD and ADHD, and degenerative disorders such as Parkinson's, Alzheimer's: movement disorders, including ataxias, supranuclear palsy, etc.); infections, such as viral infections caused by HIV-1 or HIV-2; metabolic and cardiovascular diseases and disorders (e.g., type 2 diabetes, obesity, anorexia, hypotension, hypertension, thrombosis, myocardial infarction, cardiomyopathies, atherosclerosis, etc.); proliferative diseases and cancers (e.g., different cancers such as breast, colon, lung, etc., and hyperproliferative disorders such as psoriasis, prostate hyperplasia, etc.); hormonal disorders (e.g., male/female hormonal replacement, polycystic ovarian syndrome, alopecia, etc.); sexual dysfunction, among others.

Compounds that can traverse cell membranes and are resistant to acid hydrolysis are potentially advantageous as therapeutics as they can become highly bioavailable after being administered orally to patients. However, many of these protein inhibitors only weakly inhibit function. In addition, many inhibit a variety of protein kinases and will therefore cause multiple side effects as therapeutics for diseases.

Some indolinone compounds, however, form classes of acid resistant and membrane permeable organic molecules. WO 96/22976 (published August 1, 1996 by Ballinari *et al.*) describes hydrosoluble indolinone compounds that harbor tetralin, naphthalene, quinoline, and indole substituents fused to the oxindole ring. These bicyclic substituents are in turn substituted with polar groups including hydroxylated alkyl, phosphate, and ether substituents. U.S. Patent Application Serial Nos. 08/702,232, filed August 23, 1996, entitled "Indolinone Combinatorial Libraries and Related Products and Methods for the Treatment of Disease" by Tang *et al.* (Lyon & Lyon Docket No. 221/187) and 08/485,323, filed June 7, 1995, entitled "Benzylidene-Z-Indoline Compounds for the Treatment of Disease" by Tang *et al.* (Lyon & Lyon

5

10

15

20

25

Docket No. 223/298) and International Patent Publication WO 96/22976, published August 1, 1996 by Ballinari *et al.*, all of which are incorporated herein by reference in their entirety, including any drawings, describe indolinone chemical libraries of indolinone compounds harboring other bicyclic moieties as well as monocyclic moieties fused to the oxindole ring. Applications 08/702,232, filed August 23, 1996, entitled "Indolinone Combinatorial Libraries and Related Products and Methods for the Treatment of Disease" by Tang *et al.* (Lyon & Lyon Docket No. 221/187), 08/485,323, filed June 7, 1995, entitled "Benzylidene-Z-Indoline Compounds for the Treatment of Disease" by Tang *et al.* (Lyon & Lyon Docket No. 223/298), and WO 96/22976, published August 1, 1996 by Ballinari *et al.* teach methods of indolinone synthesis, methods of testing the biological activity of indolinone compounds in cells, and inhibition patterns of indolinone derivatives, both of which are incorporated by reference herein, including any drawings.

5

10

9NSDOCID LWG 0136473A2

Other examples of substances capable of modulating kinase activity include, but are not limited to, tyrphostins, quinazolines, quinoxolines, and quinolines. The 15 quinazolines, tyrphostins, quinolines, and quinoxolines referred to above include wellknown compounds such as those described in the literature. For example, representative publications describing quinazolines include Barker et al., EPO Publication No. 0 520 722 A1; Jones et al., U.S. Patent No. 4,447,608; Kabbe et al., U.S. Patent No. 4,757,072; Kaul and Vougioukas, U.S. Patent No. 5, 316,553; 20 Kreighbaum and Comer, U.S. Patent No. 4,343,940; Pegg and Wardleworth, EPO Publication No. 0 562 734 A1; Barker et al., Proc. of Am. Assoc. for Cancer Research 32.327 (1991); Bertino, J.R., Cancer Research 3:293-304 (1979); Bertino, J.R., Cancer Research 9(2 part 1):293-304 (1979); Curtin et al., Br. J. Cancer 53:361-368 (1986); Fernandes et al., Cancer Research 43:1117-1123 (1983); Ferris et al. J. Org. 25 Chem. 44(2):173-178; Fry et al., Science 265:1093-1095 (1994); Jackman et al., Cancer Research 51:5579-5586 (1981); Jones et al. J. Med. Chem. 29(6):1114-1118; Lee and Skibo, Biochemistry 26(23):7355-7362 (1987); Lemus et al., J. Org. Chem. 54:3511-3518 (1989); Ley and Seng, Synthesis 1975:415-522 (1975); Maxwell et al., Magnetic Resonance in Medicine 17:189-196 (1991); Mini et al., Cancer Research 30 45:325-330 (1985), Phillips and Castle, J. Heterocyclic Chem. 17(19) 1489-1596 (1980); Reece et al., Cancer Research 47(11):2996-2999 (1977); Sculier et al., Cancer Immunol. and Immunother. 23:A65 (1986); Sikora et al., Cancer Letters 23:289-295

(1984); and Sikora *et al.*, Analytical Biochem. 172:344-355 (1988), all of which are incorporated herein by reference in their entirety, including any drawings.

Quinoxaline is described in Kaul and Vougioukas, U.S. Patent No. 5,316,553, incorporated herein by reference in its entirety, including any drawings.

Quinolines are described in Dolle et al., J. Med. Chem. 37:2627-2629 (1994); MaGuire, J. Med. Chem. 37:2129-2131 (1994); Burke et al., J. Med. Chem. 36:425-432 (1993); and Burke et al. BioOrganic Med. Chem. Letters 2:1771-1774 (1992), all of which are incorporated by reference in their entirety, including any drawings.

5

10

15

20

25

30

- GOOD <WO 0136473A2 1 5

Tyrphostins are described in Allen et al., Clin. Exp. Immunol, 91:141-156 (1993); Anafi et al., Blood 82:12:3524-3529 (1993); Baker et al., J. Cell Sci. 102:543-555 (1992); Bilder et al., Amer. Physiol. Soc. pp. 6363-6143:C721-C730 (1991); Brunton et al., Proceedings of Amer. Assoc. Cancer Rsch. 33:558 (1992); Bryckaert et al., Experimental Cell Research 199:255-261 (1992); Dong et al., J. Leukocyte Biology 53:53-60 (1993); Dong et al., J. Immunol. 151(5):2717-2724 (1993); Gazit et al., J. Med. Chem. 32:2344-2352 (1989); Gazit et al., J. Med. Chem. 36:3556-3564 (1993); Kaur et al., Anti-Cancer Drugs 5:213-222 (1994); King et al., Biochem. J. 275:413-418 (1991); Kuo et al., Cancer Letters 74:197-202 (1993); Levitzki, A., The FASEB J. 6:3275-3282 (1992); Lyall et al., J. Biol. Chem. 264:14503-14509 (1989); Peterson et al., The Prostate 22:335-345 (1993); Pillemer et al., Int. J. Cancer 50:80-85 (1992); Posner et al., Molecular Pharmacology 45:673-683 (1993); Rendu et al., Biol. Pharmacology 44(5):881-888 (1992); Sauro and Thomas, Life Sciences 53:371-376 (1993); Sauro and Thomas, J. Pharm. and Experimental Therapeutics 267(3):119-1125 (1993); Wolbring et al., J. Biol. Chem. 269(36):22470-22472 (1994); and Yoneda et al., Cancer Research 51:4430-4435 (1991); all of which are incorporated herein by reference in their entirety, including any drawings.

Other compounds that could be used as modulators include oxindolinones such as those described in U.S. patent application Serial No. 08/702,232 filed August 23, 1996, incorporated herein by reference in its entirety, including any drawings.

Methods of determining the dosages of compounds to be administered to a patient and modes of administering compounds to an organism are disclosed in U.S. Application Serial No. 08/702,282, filed August 23, 1996 and International patent publication number WO 96/22976, published August 1 1996, both of which are incorporated herein by reference in their entirety, including any drawings, figures or

tables. Those skilled in the art will appreciate that such descriptions are applicable to the present invention and can be easily adapted to it.

The proper dosage depends on various factors such as the type of disease being treated, the particular composition being used and the size and physiological condition of the patient. Therapeutically effective doses for the compounds described herein can be estimated initially from cell culture and animal models. For example, a dose can be formulated in animal models to achieve a circulating concentration range that initially takes into account the IC_{50} as determined in cell culture assays. The animal model data can be used to more accurately determine useful doses in humans.

5

10

15

20

25

30

BNSD/DOID WWW.2

3136413A2 X ×

Plasma half-life and biodistribution of the drug and metabolites in the plasma, tumors and major organs can also be determined to facilitate the selection of drugs most appropriate to inhibit a disorder. Such measurements can be carried out. For example, HPLC analysis can be performed on the plasma of animals treated with the drug and the location of radiolabeled compounds can be determined using detection methods such as X-ray, CAT scan and MRI. Compounds that show potent inhibitory activity in the screening assays, but have poor pharm-acokinetic characteristics, can be optimized by altering the chemical structure and retesting. In this regard, compounds displaying good pharmacokinetic characteristics can be used as a model.

Toxicity studies can also be carried out by measuring the blood cell composition. For example, toxicity studies can be carried out in a suitable animal model as follows: 1) the compound is administered to mice (an untreated control mouse should also be used); 2) blood samples are periodically obtained via the tail vein from one mouse in each treatment group; and 3) the samples are analyzed for red and white blood cell counts, blood cell composition and the percent of lymphocytes versus polymorphonuclear cells. A comparison of results for each dosing regime with the controls indicates if toxicity is present.

At the termination of each toxicity study, further studies can be carried out by sacrificing the animals (preferably, in accordance with the American Veterinary Medical Association guidelines Report of the American Veterinary Medical Assoc. Panel on Euthanasia, Journal of American Veterinary Medical Assoc., 202:229-249, 1993). Representative animals from each treatment group can then be examined by gross necropsy for immediate evidence of metastasis, unusual illness or toxicity. Gross abnormalities in tissue are noted and tissues are examined histologically. Compounds causing a reduction in body weight or blood components are less

preferred, as are compounds having an adverse effect on major organs. In general, the greater the adverse effect the less preferred the compound.

For the treatment of cancers the expected daily dose of a hydrophobic pharmaceutical agent is between 1 to 500 mg/day, preferably 1 to 250 mg/day, and most preferably 1 to 50 mg/day. Drugs can be delivered less frequently provided plasma levels of the active moiety are sufficient to maintain therapeutic effectiveness. Plasma levels should reflect the potency of the drug. Generally, the more potent the compound the lower the plasma levels necessary to achieve efficacy.

nGPCR-x mRNA transcripts may found in many tissues, including, but not limited to, brain, peripheral blood lymphocytes, pancreas, ovary, uterus, testis, salivary gland, kidney, adrenal gland, liver, bone marrow, prostate, fetal liver, colon, muscle, and fetal brain, and may be found in many other tissues. Within the brain, nGPCR-x mRNA transcripts may be found in many tissues, including, but not limited to, frontal lobe, hypothalamus, pons, cerebellum, caudate nucleus, and medulla. Tissues and brain regions where specific nGPCR mRNA transcripts are expressed are

Odd numbered nucleotide sequences ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185 will, as detailed above, enable screening the endogenous neurotransmitters/hormones/ligands which activate, agonize, or antagonize nGPCR-x and for compounds with potential utility in treating disorders including, but not limited to, thyroid disorders (e.g. thyrootoxicosis, myxoedema); renal failure; inflammatory conditions (e.g., Chron's disease); diseases related to cell differentiation and homeostasis; rheumatoid arthritis; autoimmune disorders; movement disorders; CNS disorders (e.g., pain including migraine; stroke; psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, anxiety, generalized anxiety disorder, post-traumatic-stress disorder, depression, bipolar disorder, delirium, dementia, severe mental retardation; dyskinesias, such as Huntington's disease or Tourette's Syndrome; attention disorders including ADD and ADHD, and degenerative disorders such as Parkinson's, Alzheimer's; movement disorders, including ataxias, supranuclear palsy, etc.); infections, such as viral infections caused by HIV-1 or HIV-2; metabolic and cardiovascular diseases and disorders (e.g., type 2 diabetes, obesity, anorexia, hypotension, hypertension, thrombosis, myocardial infarction, cardiomyopathies, atherosclerosis, etc.); proliferative diseases and cancers (e.g., different cancers such as breast, colon, lung,

5

10

15

20

25

30

identified in the Examples, below.

etc., and hyperproliferative disorders such as psoriasis, prostate hyperplasia, etc.); hormonal disorders (e.g., male/female hormonal replacement, polycystic ovarian syndrome, alopecia, etc.); sexual dysfunction, among others.

5

10

15

20

25

30

BNSECC : W/1 1136473A3

For example, nGPCR-x may be useful in the treatment of respiratory ailments such as asthma, where T cells are implicated by the disease. Contraction of airway smooth muscle is stimulated by thrombin. Cicala *et al* (1999) Br J Pharmacol 126:478-484. Additionally, in bronchiolitis obliterans, it has been noted that activation of thrombin receptors may be deleterious. Hauck *et al*.(1999) Am J Physiol 277:L22-L29. Furthermore, mast cells have also been shown to have thrombin receptors. Cirino *et al* (1996) J Exp Med 183:821-827. nGPCR-x may also be useful in remodeling of airway structure s in chronic pulmonary inflammation via stimulation of fibroblast procollagen synthesis. See, e.g., Chambers *et al*. (1998) Biochem J 333:121-127; Trejo *et al*. (1996) J Biol Chem 271:21536-21541.

In another example, increased release of sCD40L and expression of CD40L by T cells after activation of thrombin receptors suggests that nGPCR-x may be useful in the treatment of unstable angina due to the role of T cells and inflammation. See Aukrust *et al.* (1999) Circulation 100:614-620.

A further example is the treatment of inflammatory diseases, such as psoriasis, inflammatory bowel disease, multiple sclerosis, rheumatoid arthritis, and thyroiditis. Due to the tissue expression profile of nGPCR-x, inhibition of thrombin receptors may be beneficial for these diseases. See, *e.g.*, Morris *et al.* (1996) Ann Rheum Dis 55:841-843. In addition to T cells, NK cells and monocytes are also critical cell types which contribute to the pathogenesis of these diseases. See, e.g., Naldini & Carney (1996) Cell Immunol 172:35-42; Hoffman & Cooper (1995) Blood Cells Mol Dis 21:156-167; Colotta *et al.* (1994) Am J Pathol 144:975-985.

Expression of nGPCR-x in bone marrow and spleen may suggest that it may play a role in the proliferation of hematopoietic progenitor cells. See DiCuccio *et al.* (1996) Exp Hematol 24:914-918.

As another example, nGPCR-x may be useful in the treatment of acute and/or traumatic brain injury. Astrocytes have been demonstrated to express thrombin receptors. Activation of thrombin receptors may be involved in astrogliosis following brain injury. Therefore, inhibition of receptor activity may be beneficial for limiting neuroinflammation. Scar formation mediated by astrocytes may also be limited by inhibiting thrombin receptors. See, *e.g.*, Pindon *et al.* (1998) Eur J Biochem 255:766-

774; Ubl & Reiser. (1997) Glia 21:361-369; Grabham & Cunningham (1995) J Neurochem 64:583-591.

nGPCR-x receptor activation may mediate neuronal and astrocyte apoptosis and prevention of neurite outgrowth. Inhibition would be beneficial in both chronic and acute brain injury. See, *e.g.*, Donovan *et al.* (1997) J Neurosci 17:5316-5326; Turgeon *et al.* (1998) J Neurosci 18:6882-6891; Smith-Swintosky *et al.* (1997) J Neurochem 69:1890-1896; Gill *et al.* (1998) Brain Res 797:321-327; Suidan *et al.* (1996) Semin Thromb Hemost 22:125-133.

The attached Sequence Listing contains the sequences of the polynucleotides and polypeptides of the invention and is incorporated herein by reference in its entirety.

As described above and in Example 4 below, the genes encoding nGPCR-1 (nucleic acid sequence SEQ ID NO: 1, SEQ ID NO: 73, amino acid sequence SEO ID NO: 2, SEQ ID NO:74), nGPCR-9 (nucleic acid sequence SEQ ID NO:9, SEQ ID NO:77, amino acid sequence SEQ ID NO:10, SEO ID NO:78), nGPCR-11 (nucleic acid sequence SEQ ID NO:11, SEQ ID NO:79, amino acid sequence SEQ ID NO:12, SEO ID NO:80), nGPCR-16 (nucleic acid sequence SEQ ID NO: 21, SEQ ID NO:81, amino acid sequence SEQ ID NO: 22, SEQ ID NO:82), nGPCR-40 (nucleic acid sequence SEQ ID NO:53, SEQ ID NO:83, amino acid sequence SEQ ID NO:54, SEQ ID NO:84), nGPCR-54 (nucleic acid sequence SEQ ID NO:59, SEQ ID NO:85, amino acid sequence SEQ ID NO:60, SEQ ID NO: 86), nGPCR-56 (nucleic acid sequence SEQ ID NO:63, SEQ ID NO:87, SEQ ID NO:89, amino acid sequence SEQ ID NO:64, SEQ ID NO: 88, SEQ ID NO:90), nGPCR-58 (nucleic acid sequence SEO ID NO:3, SEQ ID NO:185, amino acid sequence SEQ ID NO:4, SEQ ID NO: 186) have been detected in brain tissue indicating that these n-GPCR-x proteins are neuroreceptors. The identification of modulators such as agonists and antagonists is therefore useful for the identification of compounds useful to treat neurological diseases and disorders. Such neurological diseases and disorders, including but are not limited to, schizophrenia, affective disorders, ADHD/ADD (i.e., Attention Deficit-Hyperactivity Disorder/Attention Deficit Disorder), and neural disorders such as Alzheimer's disease, Parkinson's disease, migraine, and senile dementia as well as depression, anxiety, bipolar disease, epilepsy, neuritis, neurasthenia, neuropathy, neuroses, and the like.

5

10

15

20

25

Methods of Screening Human Subjects

5

10

15

20

25

30

BNSCIDO: FIWE

1136473A2 1

Thus in yet another embodiment, the invention provides genetic screening procedures that entail analyzing a person's genome -- in particular their alleles for GPCRs of the invention -- to determine whether the individual possesses a genetic characteristic found in other individuals that are considered to be afflicted with, or at risk for, developing a mental disorder or disease of the brain that is suspected of having a hereditary component. For example, in one embodiment, the invention provides a method for determining a potential for developing a disorder affecting the brain in a human subject comprising the steps of analyzing the coding sequence of one or more GPCR genes from the human subject; and determining development potential for the disorder in said human subject from the analyzing step.

More particularly, the invention provides a method of screening a human subject to diagnose a disorder affecting the brain or genetic predisposition therefor, comprising the steps of: (a) assaying nucleic acid of a human subject to determine a presence or an absence of a mutation altering the amino acid sequence, expression, or biological activity of at least one seven transmembrane receptor that is expressed in the brain, wherein the seven transmembrane receptor comprises an amino acid sequence selected from the group consisting of SEQ ID NOS: 74, 186, 78, 80, 82, 84, 86, 90, and 94, or an allelic variant thereof, and wherein the nucleic acid corresponds to the gene encoding the seven transmembrane receptor; and (b) diagnosing the disorder or predisposition from the presence or absence of said mutation, wherein the presence of a mutation altering the amino acid sequence, expression, or biological activity of allele in the nucleic acid correlates with an increased risk of developing the disorder. In preferred variations, the seven transmembrane receptor is nGPCR-40 or nGPCR-54 comprising amino acid sequences set forth in SEQ ID NO: 84 for nGPCR-40 and SEQ ID NO: 86 for nGPCR-54, or an allelic variant thereof, and the disease is schizophrenia.

By "human subject" is meant any human being, human embryo, or human fetus. It will be apparent that methods of the present invention will be of particular interest to individuals that have themselves been diagnosed with a disorder affecting the brain or have relatives that have been diagnosed with a disorder affecting the brain.

By "screening for an increased risk" is meant determination of whether a genetic variation exists in the human subject that correlates with a greater likelihood

of developing a disorder affecting the brain than exists for the human population as a whole, or for a relevant racial or ethnic human sub-population to which the individual belongs. Both positive and negative determinations (i.e., determinations that a genetic predisposition marker is present or is absent) are intended to fall within the scope of screening methods of the invention. In preferred embodiments, the presence of a mutation altering the sequence or expression of at least one nGPCR-40 or nGPCR-54 seven transmembrane receptor allele in the nucleic acid is correlated with an increased risk of developing schizophrenia, whereas the absence of such a mutation is reported as a negative determination.

The "assaying" step of the invention may involve any techniques available for analyzing nucleic acid to determine its characteristics, including but not limited to well-known techniques such as single-strand conformation polymorphism analysis (SSCP) [Orita et al., Proc Natl. Acad. Sci. USA, 86: 2766-2770 (1989)]; heteroduplex analysis [White et al., Genomics, 12: 301-306 (1992)]; denaturing gradient gel electrophoresis analysis [Fischer et al., Proc. Natl. Acad. Sci. USA, 80: 1579-1583 (1983); and Riesner et al., Electrophoresis, 10: 377-389 (1989)]; DNA sequencing; RNase cleavage [Myers et al., Science, 230: 1242-1246 (1985)]; chemical cleavage of mismatch techniques [Rowley et al., Genomics, 30: 574-582 (1995); and Roberts et al., Nucl. Acids Res., 25: 3377-3378 (1997)]; restriction fragment length polymorphism analysis; single nucleotide primer extension analysis [Shumaker et al., Hum. Mutat., 7: 346-354 (1996); and Pastinen et al., Genome Res., 7: 606-614 (1997)]; 5' nuclease assays [Pease et al., Proc. Natl. Acad. Sci. USA, 91:5022-5026 (1994)]; DNA Microchip analysis [Ramsay, G., Nature Biotechnology, 16: 40-48 (1999); and Chee et al., U.S. Patent No. 5,837,832]; and ligase chain reaction [Whiteley et al., U.S. Patent No. 5,521,065]. [See generally, Schafer and Hawkins, Nature Biotechnology, 16: 33-39 (1998).] All of the foregoing documents are hereby incorporated by reference in their entirety.

Thus, in one preferred embodiment involving screening nGPCR-40 or nGPCR-54 sequences, for example, the assaying step comprises at least one procedure selected from the group consisting of: (a) determining a nucleotide sequence of at least one codon of at least one nGPCR-40 or nGPCR-54 allele of the human subject; (b) performing a hybridization assay to determine whether nucleic acid from the human subject has a nucleotide sequence identical to or different from one or more reference sequences; (c) performing a polynucleotide migration assay to

5

10

15

20

25

determine whether nucleic acid from the human subject has a nucleotide sequence identical to or different from one or more reference sequences; and (d) performing a restriction endonuclease digestion to determine whether nucleic acid from the human subject has a nucleotide sequence identical to or different from one or more reference sequences.

5

10

15

20

25

30

BNSEGG-01 (WC - 11364113A2 11 s

In a highly preferred embodiment, the assaying involves sequencing of nucleic acid to determine nucleotide sequence thereof, using any available sequencing technique. [See, e.g., Sanger et al., Proc. Natl. Acad. Sci. (USA), 74: 5463-5467 (1977) (dideoxy chain termination method); Mirzabekov, TIBTECH, 12: 27-32 (1994) (sequencing by hybridization); Drmanac et al., Nature Biotechnology, 16: 54-58 (1998); U.S. Patent No. 5,202,231; and Science, 260: 1649-1652 (1993) (sequencing by hybridization); Kieleczawa et al., Science, 258: 1787-1791 (1992) (sequencing by primer walking); (Douglas et al., Biotechniques, 14: 824-828 (1993) (Direct sequencing of PCR products); and Akane et al., Biotechniques 16: 238-241 (1994); Maxam and Gilbert, Meth. Enzymol., 65: 499-560 (1977) (chemical termination sequencing), all incorporated herein by reference.] The analysis may entail sequencing of the entire nGPCR gene genomic DNA sequence, or portions thereof; or sequencing of the entire seven transmembrane receptor coding sequence or portions thereof. In some circumstances, the analysis may involve a determination of whether an individual possesses a particular allelic variant, in which case sequencing of only a small portion of nucleic acid -- enough to determine the sequence of a particular codon characterizing the allelic variant -- is sufficient. This approach is appropriate, for example, when assaying to determine whether one family member inherited the same allelic variant that has been previously characterized for another family member, or, more generally, whether a person's genome contains an allelic variant that has been previously characterized and correlated with a mental disorder having a heritable component.

In another highly preferred embodiment, the assaying step comprises performing a hybridization assay to determine whether nucleic acid from the human subject has a nucleotide sequence identical to or different from one or more reference sequences. In a preferred embodiment, the hybridization involves a determination of whether nucleic acid derived from the human subject will hybridize with one or more oligonucleotides, wherein the oligonucleotides have nucleotide sequences that correspond identically to a portion of the GPCR gene sequence taught herein, such as

the nGPCR-40 or nGPCR-54 coding sequence set forth in SEQ ID NOS: 83 for nGPCR-40 or 85 for nGPCR-54, or that correspond identically except for one mismatch. The hybridization conditions are selected to differentiate between perfect sequence complementarity and imperfect matches differing by one or more bases. Such hybridization experiments thereby can provide single nucleotide polymorphism sequence information about the nucleic acid from the human subject, by virtue of knowing the sequences of the oligonucleotides used in the experiments.

Several of the techniques outlined above involve an analysis wherein one performs a polynucleotide migration assay, e.g., on a polyacrylamide electrophoresis gel (or in a capillary electrophoresis system), under denaturing or non-denaturing conditions. Nucleic acid derived from the human subject is subjected to gel electrophoresis, usually adjacent to (or co-loaded with) one or more reference nucleic acids, such as reference GPCR-encoding sequences having a coding sequence identical to all or a portion of SEQ ID NOS: 83 or 85 (or identical except for one known polymorphism). The nucleic acid from the human subject and the reference sequence(s) are subjected to similar chemical or enzymatic treatments and then electrophoresed under conditions whereby the polynucleotides will show a differential migration pattern, unless they contain identical sequences. [See generally Ausubel et al. (eds.), Current Protocols in Molecular Biology, New York: John Wiley & Sons, Inc. (1987-1999); and Sambrook et al., (eds.), Molecular Cloning, A Laboratory Manual, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press (1989), both incorporated herein by reference in their entirety.]

In the context of assaying, the term "nucleic acid of a human subject" is intended to include nucleic acid obtained directly from the human subject (e.g., DNA or RNA obtained from a biological sample such as a blood, tissue, or other cell or fluid sample); and also nucleic acid derived from nucleic acid obtained directly from the human subject. By way of non-limiting examples, well known procedures exist for creating cDNA that is complementary to RNA derived from a biological sample from a human subject, and for amplifying (e.g., via polymerase chain reaction (PCR)) DNA or RNA derived from a biological sample obtained from a human subject. Any such derived polynucleotide which retains relevant nucleotide sequence information of the human subject's own DNA/RNA is intended to fall within the definition of "nucleic acid of a human subject" for the purposes of the present invention.

5

10

15

20

25

In the context of assaying, the term "mutation" includes addition, deletion, and/or substitution of one or more nucleotides in the GPCR gene sequence (e.g., as compared to the seven transmembrane receptor-encoding sequences set forth of SEQ ID NOS: 74, 186, 78, 80, 82, 84, 86, 90, and 94) and other polymorphisms that occur in introns (where introns exist) and that are identifiable via sequencing, restriction fragment length polymorphism, or other techniques. The various activity examples provided herein permit determination of whether a mutation modulates activity of the relevant receptor in the presence or absence of various test substances.

5

10

15

20

25

30

5NSC-0010 kW2 - 3136473A2 L >

In a related embodiment, the invention provides methods of screening a person's genotype with respect to GPCR's of the invention, and correlating such genotypes with diagnoses for disease or with predisposition for disease (for genetic counseling). For example, the invention provides a method of screening for an nGPCR-40 or nGPCR-54 hereditary schizophrenia genotype in a human patient, comprising the steps of: (a) providing a biological sample comprising nucleic acid from the patient, the nucleic acid including sequences corresponding to said patient's nGPCR-40 or nGPCR-54 alleles; (b) analyzing the nucleic acid for the presence of a mutation or mutations; (c) determining an nGPCR-40 or nGPCR-54 genotype from the analyzing step; and (d) correlating the presence of a mutation in an nGPCR-40 or nGPCR-54 allele with a hereditary schizophrenia genotype. In a preferred embodiment, the biological sample is a cell sample containing human cells that contain genomic DNA of the human subject. The analyzing can be performed analogously to the assaying described in preceding paragraphs. For example, the analyzing comprises sequencing a portion of the nucleic acid (e.g., DNA or RNA), the portion comprising at least one codon of the nGPCR-40 or nGPCR-54 alleles.

Although more time consuming and expensive than methods involving nucleic acid analysis, the invention also may be practiced by assaying protein of a human subject to determine the presence or absence of an amino acid sequence variation in GPCR protein from the human subject. Such protein analyses may be performed, e.g., by fragmenting GPCR protein via chemical or enzymatic methods and sequencing the resultant peptides; or by Western analyses using an antibody having specificity for a particular allelic variant of the GPCR.

The invention also provides materials that are useful for performing methods of the invention. For example, the present invention provides oligonucleotides useful as probes in the many analyzing techniques described above. In general, such

oligonucleotide probes comprise 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 nucleotides that have a sequence that is identical, or exactly complementary, to a portion of a human GPCR gene sequence taught herein (or allelic variant thereof), or that is identical or exactly complementary except for one nucleotide substitution. In a preferred embodiment, the oligonucleotides have a sequence that corresponds in the foregoing manner to a human GPCR coding sequence taught herein, and in particular, the coding sequences set forth in SEQ ID NO: 83 and 85. In one variation, an oligonucleotide probe of the invention is purified and isolated. In another variation, the oligonucleotide probe is labeled, e.g., with a radioisotope, chromophore, or fluorophore. In yet another variation, the probe is covalently attached to a solid support. [See generally Ausubel et al. And Sambrook et al., supra.]

5

10

15

20

25

30

-5DOCID 4W5 0136473A2 1 >

In a related embodiment, the invention provides kits comprising reagents that are useful for practicing methods of the invention. For example, the invention provides a kit for screening a human subject to diagnose schizophrenia or a genetic predisposition therefor, comprising, in association: (a) an oligonucleotide useful as a probe for identifying polymorphisms in a human nGPCR-40 or nGPCR-54 seven transmembrane receptor gene, the oligonucleotide comprising 6-50 nucleotides that have a sequence that is identical or exactly complementary to a portion of a human nGPCR-40 or nGPCR-54 gene sequence or nGPCR-40 or nGPCR-54 coding sequence, except for one sequence difference selected from the group consisting of a nucleotide addition, a nucleotide deletion, or nucleotide substitution; and (b) a media packaged with the oligonucleotide containing information identifying polymorphisms identifyable with the probe that correlate with schizophrenia or a genetic predisposition therefor. Exemplary information-containing media include printed paper package inserts or packaging labels; and magnetic and optical storage media that are readable by computers or machines used by practitioners who perform genetic screening and counseling services. The practitioner uses the information provided in the media to correlate the results of the analysis with the oligonucleotide with a diagnosis. In a preferred variation, the oligonucleotide is labeled.

In still another embodiment, the invention provides methods of identifying those allelic variants of GPCRs of the invention that correlate with mental disorders. For example, the invention provides a method of identifying a seven transmembrane

allelic variant that correlates with a mental disorder, comprising steps of: (a) providing a biological sample comprising nucleic acid from a human patient diagnosed with a mental disorder, or from the patient's genetic progenitors or progeny: (b) analyzing the nucleic acid for the presence of a mutation or mutations in at least one seven transmembrane receptor that is expressed in the brain, wherein the at least one seven transmembrane receptor comprises an amino acid sequence selected from the group consisting of SEQ ID NOS: 74, 186, 78, 80, 82, 84, 86, 90, and 94 or an allelic variant thereof, and wherein the nucleic acid includes sequence corresponding to the gene or genes encoding the at least one seven transmembrane receptor; (c) determining a genotype for the patient for the at least one seven transmembrane receptor from said analyzing step; and (d) identifying an allelic variant that correlates with the mental disorder from the determining step. To expedite this process, it may be desirable to perform linkage studies in the patients (and possibly their families) to correlate chromosomal markers with disease states. The chromosomal localization data provided herein facilitates identifying an involved GPCR with a chromosomal marker.

The foregoing method can be performed to correlate GPCR's of the invention to a number of disorders having hereditary components that are causative or that predispose persons to the disorder. For example, in one preferred variation, the disorder is schizophrenia, and the at least one seven transmembrane receptor comprises nGPCR-40 having an amino acid sequence set forth in SEQ ID NO: 84 or an allelic variant thereof.

Also contemplated as part of the invention are polynucleotides that comprise the allelic variant sequences identified by such methods, and polypeptides encoded by the allelic variant sequences, and oligonucleotide and oligopeptide fragments therof that embody the mutations that have been identified. Such materials are useful in *in vitro* cell-free and cell-based assays for identifying lead compounds and therapeutics for treatment of the disorders. For example, the variants are used in activity assays, binding assays, and assays to screen for activity modulators described herein. In one preferred embodiment, the invention provides a purified and isolated polynucleotide comprising a nucleotide sequence encoding a nGPCR-40 or nGPCR-54 receptor allelic variant identified according to the methods described above; and an oligonucleotide that comprises the sequences that differentiate the allelic variant from the nGPCR-40 or nGPCR-54 sequences set forth in SEQ ID NOS: 83 and 88. The

5

10

15

20

25

invention also provides a vector comprising the polynucleotide (preferably an expression vector); and a host cell transformed or transfected with the polynucleotide or vector. The invention also provides an isolated cell line that is expressing the allelic variant GPCR polypeptide; purified cell membranes from such cells; purified polypeptide; and synthetic peptides that embody the allelic variation amino acid sequence. In one particular embodiment, the invention provides a purified polynucleotide comprising a nucleotide sequence encoding a nGPCR-40 seven transmembrane receptor protein of a human that is affected with schizophrenia; wherein said polynucleotide hybridizes to the complement of SEQ ID NO: 83 under the following hybridization conditions: (a) hybridization for 16 hours at 42 °C in a hybridization solution comprising 50% formamide, 1% SDS, 1 M NaC1, 10% dextran sulfate and (b) washing 2 times for 30 minutes at 60 °C in a wash solution comprising 0.1x SSC and 1% SDS; and wherein the polynucleotide encodes a nGPCR-40 amino acid sequence that differs from SEQ ID NO: 84 by at least one residue.

An examplary assay for using the allelic variants is a method for identifying a modulator of nGPCR-x biological activity, comprising the steps of: (a) contacting a cell expressing the allelic variant in the presence and in the absence of a putative modulator compound; (b) measuring nGPCR-x biological activity in the cell; and (c) identifying a putative modulator compound in view of decreased or increased nGPCR-x biological activity in the presence versus absence of the putative modulator.

Additional features of the invention will be apparent from the following Examples. Examples 1, 2, 4, 11, 12, and 13 are actual, while the remaining Examples are prophetic. Additional features and variations of the invention will be apparent to those skilled in the art from the entirety of this application, including the detailed description, and all such features are intended as aspects of the invention. Likewise, features of the invention described herein can be re-combined into additional embodiments that also are intended as aspects of the invention, irrespective of whether the combination of features is specifically mentioned above as an aspect or embodiment of the invention. Also, only such limitations which are described herein as critical to the invention should be viewed as such; variations of the invention lacking limitations which have not been described herein as critical are intended as aspects of the invention.

10

15

20

25

EXAMPLES

5

10

15

20

25

30

EXAMPLE 1: IDENTIFICATION OF nGPCR-X

A. Database search

The Celera database was searched using known GPCR receptors as query sequences to find patterns suggestive of novel G protein-coupled receptors. Positive hits were further analyzed with the GCG program BLAST to determine which ones were the most likely candidates to encode G protein-coupled receptors, using the standard (default) alignment produced by BLAST as a guide.

Briefly, the BLAST algorithm, which stands for Basic Local Alignment Search Tool is suitable for determining sequence similarity (Altschul et al., J. Molec. Biol., 1990, 215, 403-410, which is incorporated herein by reference in its entirety). Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). This algorithm involves first identifying high scoring sequence pair (HSPs) by identifying short words of length W in the query sequence that either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul et al., supra). These initial neighborhood word hits act as seeds for initiating searches to find HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Extension for the word hits in each direction are halted when: 1) the cumulative alignment score falls off by the quantity X from its maximum achieved value; 2) the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or 3) the end of either sequence is reached. The Blast algorithm parameters W, T and X determine the sensitivity and speed of the alignment. The Blast program uses as defaults a word length (W) of 11, the BLOSUM62 scoring matrix (see Henikoff et al., Proc. Natl. Acad. Sci. USA, 1992, 89, 10915-10919, which is incorporated herein by reference in its entirety) alignments (B) of 50, expectation (E) of 10, M=5, N=4, and a comparison of both strands.

The BLAST algorithm (Karlin *et al.*, Proc. Natl. Acad. Sci. USA, 1993, 90, 5873-5787, which is incorporated herein by reference in its entirety) and Gapped BLAST perform a statistical analysis of the similarity between two sequences. One measure of similarity provided by the BLAST algorithm is the smallest sum

probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a GPCR gene or cDNA if the smallest sum probability in comparison of the test nucleic acid to a GPCR nucleic acid is less than about 1, preferably less than about 0.1, more preferably less than about 0.01, and most preferably less than about 0.001.

Homology searches were performed with the program BLAST version 2.08. A collection of 340 query amino acid sequences derived from GPCR's was used to search the genomic DNA sequence using TBLASTN and alignments with an E-value lower than 0.01 were collected from each BLAST search. The amino acid sequences have been edited to remove regions in the sequence that produce non-significant alignments with proteins that are not related to GPCR's.

Multiple query sequences may have a significant alignment to the same genomic region, although each alignment may not cover exactly the same DNA region. A procedure is used to determine the region of maximum common overlap between the alignments from several query sequences. This region is called the consensus DNA region. The procedure for determining this consensus involves the automatic parsing of the BLAST output files using the program MSPcrunch to produce a tabular report. From this tabular report the start and end of each alignment in the genomic DNA is extracted. This information was used by a PERL script to derive the maximum common overlap. These regions were reported in the form of a unique sequence identifier, a start and the end position in the sequence. The sequences defined by these regions were extracted from the original genomic sequence file using the program fetchdb.

The consensus regions were assembled into a non-redundant set by using the program phrap. After assembly with phrap a set of contigs and singletons was defined as candidate DNA regions coding for nGPCR-x. These sequences were then submitted for further sequence analysis.

Further sequence analysis involved the removal of sequences previously isolated and removal of sequences related to olfactory GPCRs. The transmembrane regions for the sequences that remained were determined using a FORTRAN computer program called "tmtrest.all" [Parodi *et al.*, Comput.Appl.Biosci. 5:527-535(1994)]. Only sequences that contained transmembrane regions in a pattern found in GPCRs were retained.

5

10

15

20

25

PCT/US00/31581 WO 01/36473

cDNAs were sequenced directly using an ABI377 fluorescence-based sequencer (Perkin-Elmer/Applied Biosystems Division, PE/ABD, Foster City, CA) and the ABI PRISMTM Ready Dye-Deoxy Terminator kit with Tag FSTM polymerase. Each ABI cycle sequencing reaction contained about 0.5 μ g of plasmid DNA. Cycle-sequencing was performed using an initial denaturation at 98°C for 1 minute, followed by 50 cycles using the following parameters: 98°C for 30 seconds, annealing at 50°C for 30 seconds, and extension at 60°C for 4 minutes. Temperature cycles and times were controlled by a Perkin-Elmer 9600 thermocycler. Extension products were purified using CentriflexTM gel filtration cartridges (Advanced Genetic Technologies Corp., Gaithersburg, MD). Each reaction product was loaded by pipette onto the column, which is then centrifuged in a swinging bucket centrifuge (Sorvall model RT6000B tabletop centrifuge) at 1500 x g for 4 minutes at room temperature. Column-purified samples were dried under vacuum for about 40 minutes and then dissolved in 5 μ l of a DNA loading solution (83% deionized formamide, 8.3 mM EDTA, and 1.6 mg/ml Blue Dextran). The samples were then heated to 90°C for 15 three minutes and loaded into the gel sample wells for sequence analysis using the ABI377 sequencer. Sequence analysis was performed by importing ABI377 files into the Sequencer program (Gene Codes, Ann Arbor, MI). Generally, sequence reads of 700 bp were obtained. Potential sequencing errors were minimized by obtaining sequence information from both DNA strands and by re-sequencing difficult areas 20 using primers annealing at different locations until all sequencing ambiguities were

The following Table 5 contains the sequences of the polynucleotides and polypeptides of the invention. Start and stop codons within the polynucleotide sequence are identified by boldface type. The transmembrane domains within the polypeptide sequence are identified by underlining.

25

5

10

removed.

Table 5

The following DNA sequence beGPCR-seq1 <SEQ ID NO. 1> was identified in H. sapiens:

GTCTGGGGGTGGGGGATGCTGGGACAGGGGTCAATTGCCTGAAGCAAGTGCTCTCATCCCCCTAGCTCCTGC TGATCTAGTTGGGGCTCCAGAGTGGGGAGGAGAAAGGCACTTTGAAACTTCTCTGCCCTTACCGTCTTAGCC ATCAAACTCTGAGCTGGAGATAGTGACGATGTGACAGGAACTTTCCCTGGGCCTCTCTGGGCCACAATTCCT GGCCGAGAGAAAGAGGAGGAATGAGGTGAGCACCTTCTTCACTCCTAGGGCCATGTGGTAGAGCTGCAGTCG CACCTCCTTCTGCCAATAGGCATAGATGAGTGGGTTGAGCAGGGAGTTGCCCACGCCGAGCAGCCACAGGTA $\tt CCAGGATAGAGCAAAGCTCCCAATGAGAACAGACACAGTACGGAGAGCTTTGAAGTCGCTGGGAGTCCGTGG$ TTGAGCATGTCGCAGTAGAAGAAGACAAAGAGGGGCATGGCTGGGAAGAAGCCAACGCAGGAGAGGGTCAGC ${\tt ACGAAGTGAGGGTGAAATACAGCAAAGAAGCTGCACTGCCCTTTGTAGGCAGTCTGCTGGAACATGGGGATT}$ CCGAGTGGGAGGAAGCCAATGAGGTAAGACACTAACCACAGCCCGGCCAATGCAGGCCCCGGCCACGAACCCA $\tt CTCATGATCTTCAAGTAGCGGAAGGGCTGCTTGATGGCAAGGTACCTGTCAAAGGTGATCAGCATGACCGTG$ AGGACAGAGGCAGCTGCGGAGGAAGTGACAAATGCCATCCGCAGGCTGCACAGGGTCTTCTGTGTGGGCCGAGAAGGGCTGGAGAGCTGGTCTGTGAGTAGGCCAGAGATGGCCACACCAATCAAGGTGTCAGCCACAGCCAGA TTCAAGGTGAAGCAGAGACTGACACCATCATTCTTGTGGATCAACAGCAGCACAGCCACAGCCACTAGTGTG TTAGTAGCAATGATGAGGGAGGCCAGGACAGCAAGGATCACTCCAAATGAGAAAGATGATTCCATGTCTCGA AGTGGCAGGACTTCACTTACCAGGGCATG

The following amino acid sequence <SEQ ID NO. 2> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 1: $\frac{1}{2}$

MESSFSFGVILAVLASLIIATNTLVAVAVLLLIHKNDGVSLCFTLNLAVADTLIGVAISGLLTDQLSSPSRPT QKTLCSLEMMAFVTSSAASVLTVMLITFDRYLAIKQPFRYLKIMSGFVAGACIAGLWLVSYLIGFLPLGIPMF QQTAYKQCSFFAVFHPHFVLTLSCVGFFPAMLLFVFFYCDMLKIASMHSQQIRKMEHAGAMAGGYRSPRTPS DFKALRTVSVLIGSFALSWTPFLITGIVQVACQECHLYLVLERYLWLLGVGNSLLNPLIYAYWQKEVRLQLYH MALGVKKVLTSFLLFLSARNCGPERPRESSCHIVTISSSEFDG

The following DNA sequence beGPCR-seq3<SEQ ID NO. 3> was identified in H. sapiens:

The following amino acid sequence <SEQ ID NO. 4> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 3:

SAMGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDMPFTLLGVMRGR TPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRPRYAGLLLGCAWGQSLAFSGAALGC SWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVGFVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALA

The following DNA sequence beGPCR-seq4 <SEQ ID NO. 5> was identified in H. sapiens:

The following amino acid sequence <SEQ ID NO. 6> is the predicted amino

acid sequence derived from the DNA sequence of SEQ ID NO. 5:

 ${\tt CAGVISIPLYIPHTLFEWDFGKEIC} \underline{{\tt VFWLTTDYLLCTASVYNIVLISY}} {\tt DRYLSVSNAVSRTHFIPLR} \\ {\tt RLCKCI}$

The following DNA sequence beGPCR-seq5 <SEQ ID NO. 7> was identified in H. sapiens:

GACGTCGAAGCAGGTGATGATGCCCAGGGCGTGCACCGGGTAGGTGAGATCGGTGCGCCAGCGGGGACAGG GCGGTCAGGAGCAGCAGCCAGGTCCCTGCACACGCGGCCACCGCGTAACGACGCGGCGCCAGCGCTTGGAGC TGAGCGGGTACAGGATCCCCAGGAAGCGCTCCACGCTGATACAAGGTCATGGTGAGGATGCTGGAATACATGTT TGCGTAAAAGGCCACCGTCACCACGTCGCAAAGCACCCCGAATACCCAGTGGTGGCGGTTGCAATGGTAG TAGATTTGGAAAGGCAACACGCTGGCCAGCATCAGGTCGTGACGCTCAGGTTGATCATGAAGATGACCGACG GGGATCTGGGCCCCATGCGCCACAGCACCACAGAGAGAAGAGGTTGCCCGGGATGCTGACCGCCGCAC CAGCGAGTACACCACGGGCAGGGCCACCGCGATCGCCGGGTTCCCCCAGCATCTGCAGCGTCGCGTTGTC

The following amino acid sequence <SEQ ID NO. 8> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 7:

DNATLQMLRNPAIAVALPVVYSLVAAVSIPGNLFSLWVLCRRMGPRSPSVIFMINLSVTDLMLASVLPFQIYY HCNRHHWVFGVLCNLVVTVAFYANMYSSILTMTCISVERFLGILYPLSSKRWRRRYAVAACAGTWLLLLTAL SPLARTDLTYPVHALGIITCFDV

The following DNA sequence beGPCR-seq9<SEQ ID NO. 9> was identified in H. sapiens:

The following amino acid sequence <SEQ ID NO. 10> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 9:

PMFLLLGSLTLSDLLAGAAYAANILLSGPLTLKLSPALWFAREGGVFVALTASVLSLLGIALERSLTMARRGP APVSSRGRTLAMAAAAW

The following DNA sequence beGPCR- $\mathfrak{seq11}$ <SEQ ID NO. 11> was identified in $\mathcal{H}.$ sapiens:

The following amino acid sequence <SEQ ID NO. 12> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 11:

LLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVADFLLMICLPFRTDYYLRRRHWAFGDIPCRVGLF TLAMNRAGSIVFLTVVAADRYFKVVHPHHAVNTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCE SFIMESANGWHDIMFQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSA

RLYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYYFSSPSFPKFYNKLKICSLKPKQPGHSKTQRPEE MPIS

The following DNA sequence beGPCR-seq12<SEQ ID NO. 13> was identified in H. sapiens:

The following amino acid sequence <SEQ ID NO. 14> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 13:

WSCATTYLVNLMVADLLYVLLPFLIITYSLDDRWPFGELLCKLVHFLFYINLYGSILLLTCISVHQFLGVCHP LCSLPYRTRRHAWLGTSTTWALVVLQLLPTLAFSHTDYINGQMIWYDMTSQENFDRLFAYGIVLTLSGFLSLL GHFGVLFTDGQEPDQARGEPHEDR

The following DNA sequence beGPCR- $\mathfrak{seq14}<$ SEQ ID NO. 15> was identified in H. sapiens:

The following amino acid sequence <SEQ ID NO. 16> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 15:

RVRLVFLGVILVVAVAGNTTVLCRLXXXXXXXXXKRRKMDFLLVQLALADLYACGGTALSQLAWELLGEPRA ATGDLACRFLQLLQASGRGASAHLVVLIALERRRAVRLPHGRPLPARALAALGWLLALLLARGSGFVVRYXXX XXXXXXXTSLQPGAPLSARAWPGMRRCHWIFALLQRWHVQVYAFYEAVAGFVAPVKIMGVACGHLLSVWWRH RLKAPAGAAAWSASPGGARAPSAMPRAKVQSLKMSQLLGLLFVGCELPFADRLEAAWSSGPAGEWEGEALSAC CAWW

The following DNA sequence beGPCR-seq15<SEQ ID NO. 17> was identified in H. sapiens:

ACTCCTCGGTGCTGTTCAGGTGTTTCTGGAATGGATCTTCTAGTTTCTGCTGGTAGATCCAGGAAGCATTCTGAAGTTTTTCCATCCCTGA

The following amino acid sequence <SEQ ID NO. 18> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 17:

SGMEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGNVLVCLVILQHQ AMKTPNTYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVA ILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVT SFLFYLLPMTVISVLYYLMALRVSIAGVAG

The following DNA sequence beGPCR-seq18 <SEQ ID NO. 19> was identified in H. sapiens:

The following amino acid sequence <SEQ ID NO. 20> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 19:

 ${\tt IKMIFAIVQIIGFSNSICNPIVY} {\tt AFMNENFKKNVLSAVCYCIVNKTFSPAQRHGNSGITMMRKKAKFSLRENP}$

The following DNA sequence beGPCR-seq16 <SEQ ID NO. 21> was identified in H. sapiens:

The following amino acid sequence <SEQ ID NO. 22> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 21:

VSYSGAFSPPGDFPSMPGHNTSRNSSCDPIVTPHLISLYFIVLIGGLVGVISILFLLVKMNTRSVTTMAVINL VVVHSVFLLTVPFRLTYLIKKTWMFGLPFCKFVSAMLHIHMYLTVPILCGDPGHQIPHLLQVQRQSGILQKTA CCG

The following DNA sequence beGPCR-seq17<SEQ ID NO. 23> was identified in H. sapiens:

The following amino acid sequence <SEQ ID NO. 24> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 23:

CEYLFESWGIRLAVWAIVLLSVLCNGLVLLTVFAGGPAPLPPVKFVVGAIAGANTLTGISCGLLASVDALTLV S

The following DNA sequence beGPCR-seq20 <SEQ ID NO. 25> was identified in H. sapiens:

AACCCCATCATCTACACGCTCACCAACCGCGACCTGCGCCACGCGCTCCTGCGCCTGCTCTGCGGACGCCCACTCCTGCGGCAGAGACCCGAGTGGCTCCCAGCAGTCGGCGAGCGCGGCTGAGGCTTCCGGGGGGCCTGC

GCCGCTGCCTGCCCCGGGCCTTGATGGGAGCTTCAGCGGCTCGGAGCGCTCATCGCCCCAGCGCGACGGGCTGGACACCAGCGGCTCCACAGGCAGCCCGGT

The following amino acid sequence <SEQ ID NO. 26> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 25:

 $\underline{\texttt{NPIIYTL}} \texttt{TNRDLRHALL} \texttt{RLVCCGRHSCGRDPSGSQQSASAAEASGGLRRCLPPGLDGSFSGSERSSPQRDGLDTSGSTGSPG}$

The following DNA sequence beGPCR-seq21 <SEQ ID NO. 27> was identified in H. sapiens:

The following amino acid sequence <SEQ ID No. 28> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID No. 27:

FRCIVHPFREKLTLRKALVTIAVIWALALLIMCPSAVTLTVTREEHHFMVDARNRSYPLYSCWEAWPEKGM RRVYTTVLFSHIYLAPLALIVVMYARIARKLCXXXXXXXXXXAADPRASRRRARVVHMLVMVALFFT

The following DNA sequence beGPCR-seq22<SEQ ID NO. 29> was identified in H. sapiens:

The following amino acid sequence <SEQ ID NO. 30> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 29:

GPMPPTLLGIRQNGHAASRRLLGMDEVKGEKQLGRMFYAITLLFLLLWSPYIVACYWRVFVKACAVPHRYLAT AVWMSFAQAAVNPIVCFLLNKDLKKCLRTHAPC

The following DNA sequence beGPCR-seq24 <SEQ ID NO. 31> was identified in H. sapiens:

The following amino acid sequence <SEQ ID No. 32> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID No. 31:

VYRVITIIKLFGRGSEWTDSDNYKDTDETFILFVLMYDLSYFLTAGAFICKMVPFVQSTAVVTEILTMTCIAV ERHQGLVHPFKMKWQYTNRRAFTMLGEATGCANGSVNDILHYRI

The following DNA sequence beGPCR- ${\tt seq27}$ <SEQ ID NO. 33> was identified in H. sapiens:

GCACATGAGCACCAGGACCGTCTTCCTGCGGCAGCGCAGCCTCTTGCGGATCTGCTCTGTAATCCAG
GGACCGCCTTGAACCAGAGCTCCCGGGAGATCCTGGCATAGCACAGGGTCATGGTGACCACGGGGCCCACG
AATTCTATGCCAAAGATAAAGAGGAAGTAGGACTTGTAGTAGAGCTGCTGGTCCACAGGCCAGATCTGGCC
GCAGAAGATCTTTTCCTGGCTCTTGACAATGACGAGGACCGTCTCGGTGGTGAAGTAGGCGGAAGGGATGG
CGATCAGGATGGACACCGTCCACACCAAGGCAATCAGGCCAGTGGCTGTTTGGCACTTCATTCGTGGTCTC
AGCGGATGGACAATAGCCAGATACCTAGGGCAAGAACACAAGTGGAGGCAGCC

The following amino acid sequence <SEQ ID NO. 34> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 33:

GCLHLCSCPRYLAIVHPLRPRMKCQTATGLIALVWTVSILIAIPSAYFTTETVLVIVKSQEKIFCGQIWPVDQ QLYYKSYFLFIFGIEFVGPVVTMTLCYARISRELWFKAVPGFQTEQIRKRLRCRRKT<u>VLVLMCILTAYVLCWA</u> PFYGFTIVRDFFPTVFVKEKHYLTAF<u>Y</u>IVECIAMSNSMINTLCFVTVKNDTVKYFKKIMLL

The following DNA sequence beGPCR-seq28 <SEQ ID NO. 35> was identified in H. sapiens:

The following amino acid sequence <SEQ ID NO. 36> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 35:

LIPVFLILFIALVGLVGNGFVLWLLGFRMRRNAFSVYVLSLAGADFLFLCFQIINCLVYLSNFFCSISINFPS FFTSVMTFAYLVGLSMLSAISTECCLSVLRPIWYCCCCPRNLSTVMCALPWALSLLLNTLEGKFCGFLVSNGD YGWCWTFDFITAVWL

The following DNA sequence beGPCR-seq31<SEQ ID NO. 37> was identified in H. sapiens:

GAGAGTCTGATTCTGACTTACATCACATATGTAGGCCTGGGCATTTCTATTTGCAGCCTGATCCTTTGCTTGT
CCGTTGAGGCTCTAGTCTGGAGCCAAGTGACAAAGACAAGAGTCACCTATTTACGCCATGTGTGCATTGTTAA
CATTGCAGCCACTTTGCTGATGGCAGATGTGTGGTTCATTGTGGCTTCCTTTCTTAGTGGCCCAATAACACAC
CACAAGGGATGTGTGGCAGCCACATTTTTTGGTCATTTCTTTACCTTTCTGTATTTTTCTGGATGCTTGCCA
AGGCACTCCTTATCCTCTATGGAATCATGATTGTTTTC

The following amino acid sequence <SEQ ID NO. 38> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 37:

 ${\tt ESLILTYITYVGLGISICSLILCLSVEVLVWSQVTKTEITYLRHVCIVNIAATLLMADVWFIVASFLSGPITH\\ HKGCVAATFFGHFFYLSVFFWMLAKALLILYGIMIVF}$

The following DNA sequence beGPCR-seq32 <SEQ ID NO. 39> was identified in H. sapiens:

TTGTGTGGCAGTAGAGAGATGTCAGGCTTCAGAGTCAACAAGAACTGGATTTCAAACTGGATTTGAGGACCCC
CACCTTTGGTAAGTGACTTATTATCTGCGAGCCTCTGTTCTCTCTTCTTTAAATGAGGACAGTAAATCCCAT
ACGCAGGGTGGTGGGGAGAATCAGAGATGATACAGCTGGTGATCACATCTGGTTTGTGTTCCCAGGGGCACC
AGACTAGGGTTTCTGAGCATGGATCCAACCGTCCCAGTCTTCGGTACCAAAACTGACACCAATCAACGGACGTG
AGGAGACTCCTTGCTACAATCAGACCCTGAGCTTCACGGTGCTGACGTGCATCATCTCCCTTGTGGACTGAC
AGGAAACGCGGTAGTGCTCTGGGCTACCGCATGCGCAGGAACGCTGTCTCCATCTACATCCTCAAC
CTGGCCGCAGCAGACTTCCTCTCTCCTCAGCTTCCAGATTAACGTTCGCCATTACGCCTCATCAATATCAGCC
ATCTCATCCGCAAAAATCCTCCTTTCTGTGATGACCTTTCCCTACTTTACAGGCCTGAGTATGCTGAGCGCAT
CAGCACCGAGCGCTGCTGTCTTCTGTGGCCCATCTGGTACC

The following amino acid sequence <SEQ ID NO. 40> is the predicted amino

acid sequence derived from the DNA sequence of SEQ ID NO. 39:

LCGSREMSGFRVNKNWISNWIGPPPLVSDLLSASLCFSLLMRTVNPIRQGGGENQRYSWSHLVCVPRGTRLGF LSMDPTVPVFGTKLTPINGREETPCYNQTLSFTVLTCIISLVGLTGNAVVLWLLGYRMRRNAVSIYILNLAAA DFLFLSFQIIRSPLRLINISHLIRKILVSVMTFPYFTGLSMLSAISTERCLSVLWPIWY

The following DNA sequence beGPCR-seq33 <SEQ ID NO. 41> was identified in H. sapiens:

The following amino acid sequence <SEQ ID NO. 42> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 41:

TESKATRTLGIVMGVFVLCWLPFFVLTITDPFINFTTLEDLYNVFLWLGYFNSAFNPILYGMLYPWFRKALRM IVTGMIFHPDSSTLSLFSAHAAVFIIQDSF

The following DNA sequence beGPCR-seq34<SEQ ID NO. 43> was identified in H. sapiens:

The following amino acid sequence <SEQ ID No. 44> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID No. 43:

 $LHQRGMVAKRQEMLAAFLVSWLPYLVDAVIDAYMNFITPPYVYE\\ ilvwcvyynsamnpliyaffyQwfgkaIK\\ LIVSGKVLRTDSSTTNLFSEEVETDKHYCRDLKTNLKLRSTAKINTWTRGKHDHMPSCRTIHSTVVLKHLLSS\\ CI$

The following DNA sequence beGPCR-seq35 <SEQ ID NO. 45> was identified in H. sapiens:

The following amino acid sequence <SEQ ID NO. 46> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 45:

LERGPRSILYAVLGFGAVLAAFGNLLVMIAILHFQLHTPTNFLIASLACADFLVGVTVMPFSTVRSVESCWYF GDSYCKFHTCFDTSFCFASLFHLCCISVDRYIAVTDPLTYPTKFTVSVSGICIVLSWFFSVTYSFSIFYTGAN EEGIEELVVALTCVGGCQAPLNQNWVLLCFLLFFIPNVAMVFIYSKIFLVAKHQARKIESTASQAQSFSESYK ERVAKRERKAAKTLGIAMAAFL

The following DNA sequence beGPCR-seq36 <SEQ ID NO. 47> was identified in H. sapiens:

The following amino acid sequence <SEQ ID NO. 48> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 47:

NQVALLLRP<u>LALSMAFINSCLNPVLYVFI</u>GHDFWEHLLHSLLAALERALSEEPDSAIPAPRQMSPLHDPISYS IFPPLNPLPKQLYHNPTSNRIENKPQLLSELYVLGHVLEYNLKCLEDGGKKQTRSHSLEEDSSPRLKQKKRLS CDKTSHKIGSGPAAMTLCNPEHQETAILLNQSQVWTYMSGKTQRATLILKLQGIAQCHQDPFDDL

The following DNA sequence beGPCR-seq37<SEQ ID NO. 49> was identified in H. sapiens:

The following amino acid sequence <SEQ ID NO. 50> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 49:

The following DNA sequence beGPCR-seq38 <SEQ ID NO. 51> was identified in H. sapiens:

TTACTTATTCTGCCCTTTATCCAACTTTTAATTCCTTTGCTATTCTCTGCCTCATTTTCTGGCCTCATTTT
CCCTATTATCCTGCCTCACATTGATCAAGGGATGAGGCTGGCAGGATCCGGAACCCACAGGGCCCGTGGGCC
ATGAGAGGCTCCTGGACTTGAACCTCAGGACACTCCCACTCTGGCTGCCGGCAGGGATGGAAGCTGGATGAGC
AGGCAGGAGCTGGCAGTGGGGGTGGAGAGCCATAGGCTATTGGGGTGACAGGCTTTGCGGGAGAGGTTCAAGTCA
CCCCATGGGAGCTTGGGCCCTTGGGGCCTCTTATTTCTCACCCCAGGCTTTCCCGGGAGAGGTTCAAGTCA
GAAGATGCCCCAAAGATCCACGTGGCCCTGGGTGGCAGCCTGTTCCTCGAATCTGGCCTTCTTGTCAAAGGGTCTCAAGGGTCTGAATCTGGCCTTCTGGATGCCCTGAATCTTCCTGCTCTG
TGCGGAGTGGCCTGGATGGGCCTTGAAGCCTTCCACCTCTACCTGCTCTCAGGGTCTTCAACACCTACTTC
GGGCACTACTTCCTGAAGC

The following amino acid sequence <SEQ ID NO. 52> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 51:

ETYSALYPTFNSLCYSPASFSGLIFPIILPHIDQGMRLAGSGTHRAPWAMRGSWTTSGHSHSGCRQGWKLDEQ AGAGSGGGPPAIGVDRLGCLMGAPHGSCGPLGPLISHPRLSRERFKSEDAPKIHVALGGSLFLLNLAFLVNVG SGSKGSDAACWARGAVFHYFLLCAFTWMGLEAFHLYLLAVRVFNTYFGHYFL

The following DNA sequence beGPCR-seq40 <SEQ ID NO. 53> was identified in

H. sapiens:

AATTGGTCGGAGAGTGCAGCTGCTTGAAATGGAGGATTGAAATCATCACCAGGAGGTTTCCAAACACAGCCAG CACAGCCCCAAAGCCAAACACTATGTACAGAATCACCCGGGATCCCGGCGAGAAGGGGATTTTCACACAGGAC CCATTCACGTTCGCGTAGCACAGCTGCACAGCGCACCAGCAGGGGTGAATTGCTGCTCATAACGCTGGTATTTA CATATGGAGAAATTTTGTCCTTGTTGATTATCACAAAAATACAGGATTGTTCCTGATTTTCATTGCTCCTGC GGAAAAAAACACATATTCACCAGGATGCCAGAGGAAATGATCA

The following amino acid sequence <SEQ ID NO. 54> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 53:

The following DNA sequence beGPCR-seq41 <SEQ ID NO. 55> was identified in H. sapiens:

The following amino acid sequence <SEQ ID NO. 56> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 55:

LTDFLSFFIPTFIMIILYGNIFLVARRQAKKIENTGSKTESSSESYKARVARRERKAAKTLGVTVVAFMISWL PYSIDSLIDAFMGFITPACIYEICCWCAYYNSAMNPLIYALFYPWFRKAIKVIVTGQVLKNSSATMNLFSEHI AVGTKFRIPLKLPSEMSFKSSKTMNEQINCSSNKQINVFQSCDV

The following DNA sequence nGPCR-seq53 <SEQ ID NO. 57> was identified in H. sapiens:

The following amino acid sequence <SEQ ID NO. 58> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 57:

CGKETLIPVFLILFIALVGLVGNGFVLWLLGFRMRRNAFSVYVLSLAGADFLFLCFQIINCLVYLSNFFCSIS
INFPSFFTTVMTCAYLAGLSMLSTVSTERCLSVLWPIWYRCRRPRHLSAVVCVLLWALSLLLSILEGKFCGFL
FSDGDSGWCQTFDFITAAWLIFLFMVLCGSSLALLVRILCGSRGLPLTRLYLTILLTVLVSLLCGLPFGIQWF
LILWIWKDSDVLFCHIHPVSVVLSSLNSSANPIIYFFVCSFRKQWRXQHPILKLALQRALQDTAEVDHSEGCF
RQGTRRFKEAFWCRDGPLYFHHIYVALRGNFA

The following DNA sequence nGPCR-seq54<SEQ ID NO. 59> was identified in H. sapiens:

The following amino acid sequence <SEQ ID NO. 60> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 59:

FRYCVIIHPMSCFSIHKTRCAVVACAVVWIISLVAVIPMTFLITSTNRTNRSACLDLTSSDELNTIKWYNLIL TASTFCLPLVIVTLCYTTIIHTLTHGLQTDSCLKQKARRLTILLLLAFYVCFLPFHILRVIQDRISACFQSVV PLRIRSMKLTSFLDHYAALNTFGNLLLYVVVSDNFQQAVCSTVRCK

The following DNA sequence nGPCR-seq55 <SEQ ID NO. 61> was identified in H. sapiens, where the underlined ATG identifies a probable start codon:

The following amino acid sequence <SEQ ID NO. 62> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 61:

 $\label{thm:mantigepeevs} \textbf{MANTTGEPEEVSGALSPPSASA} \underline{\textbf{VVKLVLLGLIMCVSLAGNAILSLLVLKERALHKAPYYFLLDLCLADGIRSA}\\ \underline{\textbf{VCFPFVLASVRHGSSWTFSALSCKIVAFMAVLFCFHAAFMLFCISV}} \underline{\textbf{TRYMAIAHHRFYAKRMTLWTCAAE}}$

The following DNA sequence nGPCR-seq56 <SEQ ID NO. 63> was identified in H. sapiens:

The following amino acid sequence <SEQ ID NO. 64> is the predicted amino

acid sequence derived from the DNA sequence of SEQ ID NO. 63:

REKTDQPSGMMPFCHNIINISCVKNNWSNDVRASLYSLMVLIILTTLVGNLIVIVSISHFKQLHTPTNWLIHS
MATVDFLLGCLVMPYSMVRSAEHCWYFGEVFCKIHTSTDIMLSSASIFHLSFISIDRYYAVCDPLRYKAKMNI
LVICVMIFISWSVPAVFAFGMIFLELNFKGAEEIYYKHVHCRGGCSVFFSKISGVLTFMTSFYIPGSIMLCVY
YRIYLIAKEQARLISDANQ

The following DNA sequence nGPCR-seq57<SEQ ID NO. 65> was identified in H. sapiens:

The following amino acid sequence <SEQ ID NO. 66> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 65:

YIKECFLKVPVEEALYLTSKYRLSICNLKIQNLKCSKIWNFLSINMMPQVENSTPEAFAVWFNVCKLCFMPKI
INIVQNYFQTMCIRCININKFCVTWEPFPRYIIMNVIFRNPKSKTFLVSNILGKGYSTCTTVILLLTFTPEML
KVCFSPTGVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFVVKILS
LFRVEIPGQSLLSFPSIIHRAFLRPSFDKARVDTIIHKNQYKVISLPCFIISIIKKLSSGAIQPGIIKSRSYR
ETKSEYLASIARHWFFTRSMHKTIKIYMPRFHPGL

The following DNA sequence nGPCR- $\mathfrak{seq58}$ <SEQ ID NO. 67> was identified in H. sapiens:

ACTACCATGGAAGCTGACCTGGGTGCCACTGGCCACAGGCCCCGCACAGAGCTTGATGATGAGGACTCCTACC GTGGCTGGCCGGCTCCCAGGCCCGGCATGGAGCTGGCACGCGTCTGGCGCTGCTCCTGCTCAGCCTGGCCCTC TCTGACTTCTTGTTCCTGGCAGCAGCGGCCTTCCAGATCCTAGAGATCCGGCATGGGGGACACTGGCCGCTGG GGACAGCTGCCTGCCGCTTCTACTACTTCCTATGGGGCGTGTCCTACTCCTCCGGCCTCTTCCTGCTGGCCGC CCTCAGCCTCGACCGCTGCTGCCCGTGCCCACACTGGTACCCTGGGCACCGCCCAGTCCGCCTGCCC GGTCCTGGGGGGGCTTCCTGCCTTTCCTCCTGCTGCTCGTCTGCCACGTGCTCACCCAGGCCACAGCCTGTCGC ATGTGGTCCTGAGGCTGCCCTACCAGCTGGCCCAGCTGCTCTACCTGGCCTTCCTGTGGGACGTCTACTCTGG $\tt CTACCTGCTCTGGGAGGCCCTGGTCTACTCCGACTACCTGATCCTACTCAACAGCTGCCTCAGCCCCTTCCTC$ TGCCTCATGGCCAGTGCCGACCTCCGGACCTGCTGCGCTCCTTCGTCCTTCGCGGCAGCTCTCTGCG AGGAGCGGCCGGCCACTCACCCCACTGAGCCACAGACCCAGCTAGATTCTGAGGGTCCAACTCTGCCAGA GCCGATGGCAGAGGCCCAGTCACAGATGGATCCTGTGGCCCAGCCTCAGGTGAACCCCACACTCCAGCCACGA TCGGATCCCACAGCTCAGCCACAGCTGAACCCTACGGCCCAGCCACAGTCGGATCCCACAGCCCAGCCACAGC TGAACCTCATGGCCCAGCCACAGTCAGATTCTGTGGCCCAGCCACAGGCAGACACTAACGTCCAGACCCCTGC ACCTGCTGCC

The following amino acid sequence <SEQ ID NO. 68> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 67:

TTMEADLGATGHRPRTELDDEDSYPOGGWDTVFLVALLLLGLPANGLMAWLAGSQARHGAGTRLALLL LSLALSDFLFLAAAAFOILEIRHGGHWPLGTAACRFYYFLWGVSYSSGLFLLAALSLDRCLLALCPHW WO 01/36473

YPGHRPVRLPLWVCAGVWVLATLFSVPWLVFPEAAVWWYDLVICLDFWDSEELSLRMLEVLGGFLPFL LLLVCHVLTQATACRTCHRQQQPAACRGFARVARTILSAYVVLRLPYQLAQLLYLAFLWDVYSGYLLW EALVYSDYLILLNSCLSPFLCLMASADLRTLLRSVLSSFAAALCEERPGSFTPTEPQTQLDSEGPTLP EPMAEAQSQMDPVAQPQVNPTLQPRSDPTAQPQLNPTAQPQSDPTAQPQLNLMAQPQSDSVAQPQADT NVQTPAPAA

The following DNA sequence nGPCR-seq59 <SEQ ID NO. 69> was identified in H. sapiens:

The following amino acid sequence <SEQ ID NO. 70> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 69:

YRPEHAGLHQHQALPVHPVAHLVPLPPPHTPVSSRVSCSGPCPCCRASWNGCSVASCLVVLILFGVKHQISSQ SHGFFYVWFSAGPARFCWSGSFVDPGRCPPGCTPSCSECWSSSSVTCPLAFSDSYFSGSTWICHVRLVSIFLS TLNSSANPIIYFFMGSFRQLQNRKTLLVLQRALQDTPEVEEGRWRLSEETLELSSRLGPGRASALSV

The following DNA sequence nGPCR-seq60 <SEQ ID NO. 71> was identified in H. sapiens:

The following amino acid sequence <SEQ ID NO. 72> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 71:

LSSNYYRNPFAIYLLDVACADLIFLGCHMVAIVPDLLQGRLDFPGFVQTSLATLRFFCYIVGLSLLA
AVSVEQCLAALFPAWYSCRRPRHLTTCVCALTWALCLLLHLTTCVCALTWALCLLLHLLLSGACTLL
LSGACTQFFGEPSRHLCRTLWLVAAVLLALLCCTMCGASLMLLLRVERGPQRPPPRGFPGLILLTVL
LFSSAACLRH

The following DNA sequence nGPCR-1 <SEQ ID NO. 73> was identified in H. sapiens:

CGTGGGCAACTCCTGCTCAACCCACTCATCTATGCCTATTGGCAGAAGGAGGTGCGACTGCAGCTCTACCAC ATGGCCCTAGGAGTGAAGAAGGTGCTCACCTCATTCCTCCTCTTTCTCTCGGCCAGGAATTGTGGCCCAGAGA GGCCCAGGGAAAGTTCCTGTCACATCGTCACTATCTCCAGCTCAGAGTTTGATGGCTAA

The following amino acid sequence <SEQ ID NO. 74> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 73:

MESSFSFGVILAVLASLIIATNTLVAVAVLLLIHKNDGVSLCFTLNLAVADTLIGVAISGLLTDQLSSPSRPT QKTLCSLRMAFYTSSAASVLTVMLITFDRYLAIKQPFRYLKIMSGFVAGACIAGLWLVSYLIGFLPLGIPMF QQTAYKGQCSFFAVFHPHFVLTLSCVGFFPAMLLFVFYCDMLKIASMHSQQIRKMEHAGAMAGGYRSPRTPS DFKALRTVSVLIGSFALSWTPFLITGIVQVACQECHLYLVLERYLWLLGVGNSLLNPLIYAYWQKEVRLQLY HMALGVKKVLTSFLLFLSARNCGPERPRESSCHIVTISSSEFDG

The following DNA sequence TL-GPCR-seq5 <SEQ ID NO. 75> was identified in H. sapiens.

AACTGGAAGGGCAGCCGTCTGCCGCCCACGAACACCTTCTCAAGCACTTTGAGTGACCACGGCTTGCAAGCTG GTGGCTGGCCCCCGGGTCCCGGGCTCTGAGGCACGGCCGTCGACTTAAGCGTTGCATCCTGTTACCTGGAGA CCCTCTGAGCTCTCACCTGCTACTTCTGCCGCTGCTTCTGCACAGAGCCCGGGCGAGGACCCCTCCAGGATGC AGGTCCCGAACAGCACCGGCCCGGACAACGCGACGCTGCAGATGCTGCGGAACCCGGCGATCGCGGTGGCCCT GCCCGTGGTGTACTCGCTGGTGGCGGCGGTCAGCATCCCGGGCAACCTCTTCTCTCTGTGGGTGCTGTGCCGG CGCATGGGGCCCAGATCCCCGTCGGTCATCTTCATGATCAACCTGAGCGTCACGGACCTGATGCTGGCCAGCG TGTTGCCTTTCCAAATCTACTACCATTGCAACCGCCACCACTGGGTATTCGGGGTGCTGCTTTGCAACGTGGT GACCGTGGCCTTTTACGCAAACATGTATTCCAGCATCCTCACCATGACCTGTATCAGCGTGGAGCGCTTCCTG GGGGTCCTGTACCCGCTCAGCTCCAAGCGCTGGCGCCGCCGTCGTTACGCGGTGGCCGCGTGTGCAGGGACCT GGCTGCTGCTCCTGACCGCCTGTCCCCGCTGGCGCGCACCGATCTCACCTACCCGGTGCACGCCCTGGGCAT CATCACCTGCTTCGACGTCCTCAAGTGGACGATGCTCCCCAGCGTGGCCATGTGGGCCGTGTTCCTCTTCACC ATCTTCATCCTGCTGTTCCTCATCCCGTTCGTGATCACCGTGGCTTGTTACACGGCCACCATCCTCAAGCTGT TGCGCACGGAGGAGCCGCCGGGAGCAGCGGAGCGCGCGGTGGCCCTGGCCGCGGTGGTCTTGCTGGC CTTTGTCACCTGCTTCGCCCCCAACACTTCGTGCTCCTGGCGCACATCGTGAGCCGCCTGTTCTACGGCAAG ${\tt ACTTTGCGTCCCGGGAATTCCAGCTGCGCCTGCGGGAATATTTGGGCTGCCGGGGTGCCCAGAGACACCCT}$ GGACACGCGCCGGAGAGCCTCTTCTCCGCCAGGACCACGTCCGTGCGCTCCGAGGCCGGTGCGCACCCTGAA GGGATGGAGGGAGCCACCAGGCCCGGCCTCCAGAGGCAGGAGAGTGTGTTCTGAGTCCCGGGGGCGCAGCTTG GAGAGCCGGGGGCGCAGCTTGGAGGATCCAGGGGCGCATGGAGAGGCCACGGTGCCAGAGGTTCAGGGAGAAC GGGTGCTTGTTATCCTGCAGAGGGTGCCTCTGCCTCTGTGTGAGGGGACAGCTTGTGTCACCACGCCCGGC TACCCGAGGATGGATATTCAACCAGCCCCACCGCCTACCCGACTCGGTTTCTGGATATCCTCTGTGGGCGAAC TGCGAGCCCCATTCCCAGCTCTTCTCCCTGCTGACATCGTCCCTTAGCACACCTGTCCATACCCGAGGATGGA TATTCAACCAGCCCACCGCCTACCCGACTCGGTTTCTGGATATCCTCTGTGGGCGAACTGCGAGCCCCATTC $\tt CCAGCTCTTCTCCCTGACATCGTCCCTTAGTTGTGGTTCTGGCCTTCTCCATTCTCCAGGGGTTCTG$ GTCTCCGTAGCCCGGTGCACGCCGAAATTTCTGTTTATTTCACTCAGGGGCACTGTGGTTGCTGTGGTTGGAA ACCCCCTCGTGCCGAATTC

The following amino acid sequence <SEQ ID NO. 76> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 75.

MQVPNSTGPDNATLQMLRNPAIAVALPVVYSLVAAVSIPGNLFSLWVLCRRMGPRSPSVIFMINLSVTDLMLA SVLPFQIYYHCNRHHWVFGVLLCNVVTVAFYANMYSSILTMTCISVERFLGVLYPLSSKRWRRRYAVAACAG TWLLLLTALSPLARTDLTYPVHALGIITCFDVLKWMLPSVAMWAVFLFTIFILLFLIPFVITVACYTATILK LLRTEEAHGREQRRRAVGLAAVVLLAFVTCFAPNNFVLLAHIVSRLFYGKSYYHVYKLTLCLSCLNNCLDFFV YYFASREFOLRLREYLGCRRVPRDTLDTRRESLFSARTTSVRSEAGAHPEGMEGATRPGLOROESVF

The following DNA sequence nGPCR-9 <SEQ ID NO. 77> was identified in H. sapiens:

ATGGAGTCGGGGCTGCTGCGGCCGGCGGCGGTGAGCGAGGTCATCGTCCTGCATTACAACTACACCGGCAGCTCCGCGGTGCGGCTACCAGCCGGGTGCCGCCCTCGCGCCCGACGCCGTGGTGTGCCTTGCGCGTTTGCGCCTT

The following amino acid sequence <SEQ ID NO. 78> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 77:

MESGLLRPAPVSEVIVLHYNYTGKLRGARYQPGAGLRADAVVCLAVCAFIVLENLAVLLVLGRHPRFHAPMFL LLGSLTLSDLLAGAAYAANILLSGPLTLKLSPALWFAREGGVFVALTASVLSLLAIALERSLTMARRGPAPVS SRGRTLAMAAAAWGVSLLLGLLPALGWNCLGRLDACSTVLPLYAKAYVLFCVLAFVGILAAICALYARIYCQV RANARRLPARFGTAGTTSTRARRKPRSLALLRTLSVVLLAFVACWGPLFLLLLLDVACPARTCPVLLQADFFL GLAMANSLLNPIIYTLTNRDLRHALLRLVCCGRHSCGRDPSGSQQSASAAEASGGLRRCLPPGLDGSFSGSER SSPORDGLDTSGSTGSPGAPTAARTLVSEPAAD

The following DNA sequence nGPCR-11 <SEQ ID NO. 79> was identified in H. sapiens:

The following amino acid sequence <SEQ ID NO. 80> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 79:

MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVADFLLMICLPFRTD YYLRRRHWAFGDIPCRVGLFTLAMMRAGSIVFLTVVAADRYFKVVHPHHAVNTISTRVAAGIVCTLWALVILG TVYLLLENHLCVQETAVSCESFIMESANGWHDIMFQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKAT RFIMVVAIVFITCYLPSVSARLYFLWTVPSSACDPSVHGALHITLSFTYMMSMLDPLVYYFSSPSFPKFYNKL KICSLKPKQPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSDGQWDPHIVEWH

The following DNA sequence nGPCR-16 <SEQ ID NO. 81> was identified in H. sapiens:

ATGACAGGTGACTTCCCAAGTATGCCTGGCCACAATACCTCCAGGAATTCCTCTTGCGATCCTATAGACACCCCACTTAATCAGCCTCTACTTCATAGTGCTTATTGGCGGGCTGGTGGTGGTTCATTTCCATTCTTTTCCTCCTGGTGAAAATGAACACCCGGTCAGTGACCACCATGGCGGTCATTAACTTGGTGGTGGTCCACAGCGTTTTTCTGCTGACAGTGCCATTTCGCTTGACCTCATCAAGAAGACTTGGATGTTTTGGGCTGCCCTTCTGCAAATTTGTG

AGTGCCATGCTGCACATCCACATGTACCTCACGTTCCTATTCTATGTGGTGATCCTGGTCACCAGATACCTCA TCTTCTTCAAGTGCAAAGACAAAGTGGAATTCTACAGAAAACTGCATGCTGTGGCTGCCAGTGCTGGCATGTG GACGCTGGTGATTGTCATTGTGGTACCCCTGGTTGTCTCCCGGTATGGAATCCATGAGGAATACAATGAGGAG CACTGTTTTAAATTTCACAAAGAGCTTGCTTACACATATGTGAAAATCATCAACTATATGATAGTCATTTTTG TCATAGCCGTTGCTGTGATTCTGTTGGTCTTCCAGGTCTTCATCATTATGTTGATGGTGCAGAAGCTACGCCA TTCCTTCCCTACCAGTTCTTTAGGATCTATTACTTGAATGTTGTGACGCATTCCAATGCCTGTAACAGCAAGG TTGCATTTTATAACGAAATCTTCTTGAGTGTAACAGCAATTAGCTGCTATGATTTGCTTCTCTTTTGTCTTTTGG GGGAAGCCATTGGTTTAAGCAAAAGATAATTGGCTTATGGAATTGTGTTTTTGTGCCGT**TAG**CCACAAACTACA GTATTCATATTTGCTTCCTTTATATTGGGAATAAAAATGGGTATAGGGGAGGTAAGAATGGTATTTCATTACT TGATCAAAACCATGCCTTGATGTACCCAAAACAAAGGACTATAAAATGCAAGAGCCCTCATTGTAGTCCTTA TGGGATCCCTCCCATCTCTGAGTGATGGCCGTACAAAGACCAGTGTTGTTGAATCCACCTGGAGTTGCAATAT TACATTATTTTCCAGTACAGAATGTCTGTGTGGCCCATGAAAGCAACATAGGTTTTAAGAGTTTTAGAGTTTC ATTAGCTCATTCTAAGTTCCTCTGTTTGAAGCATGGTCTCTTAGGTTTTGGACTGAACTCAGACCTTTAGGTT TTTTCATCCCACTTCACCTTAGGTAAGTAAATTCTGGCCACCAGCTCCAAAGACACAAACTCTCCTTCG CTAACCAGGTTAGATGTCCCATTCATCTCATGCCCTGATAAAAACTGATAAGGGGAGAGATAGTTAAAAATT GTTATAACAAGGGTTTCTAGATTTGTCCTGAAAGGTCGTTTAAGGACTTGGGGATCAACTTCCTCAATTAT CACCAATTGCACTGTTGCTCCAAAAATCATTTAAAAGCTTACTGGACATATCTACATAATGGTGAAACTGTAA ${\tt TTTAGAGACTATCCCTGACTAATGTGCTGGTAGGCATTAAAATGAGTTCCCAAGGGAAGTGATTAAAATTTTT}$ AGTTCTGGGGTACATGTGCAGAATGTGCAGGTTTGTTACATAGGTATACACGTGCCATGGTGGTTTGCGGCAC CTGTCAACCCATCTACATTAGGTATTTCTCCTAATGCTCTCCCCTAGCCCCCCACCCCTGGACAGGCCC CATTGTGTGATGTTCCCCTCCCTGTGTCCATGTGTTTTCATTGTTCAACTCCCACTTCTAAGTGAGAACATGC GGTGTTTGGTTTCTGTTCCTGTGTTAGTTTGCTGAGAATGATGGTTTCCAGGTTAAAATTATATTTTTTAA ATAAATGAAAACTGTGTTTTTAAAAGAGGACTTTTGAGAAGTATATAGAAAAACCATTAATTTAGACTCTGTG ATCATAATCCTTTAAAATATAGGAAAAATAACTAATGGGAACTAGGCTTAATACTCGGGATGAAATAATCTGT ACAACAAACTCCCATGACACATGTTTACCTATGTAACAAACCTGCACATGTACCCCTGAACTTAAAAATAAAAT TTAAAGTATAATAAAAATAATATGGATTTTCTTT

The following amino acid sequence <SEQ ID NO. 82> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 81:

MTGDFPSMPGHNTSRNSSCDPIVTPHLISLYFIVLIGGLVGVISILFLLVKMNTRSVTTMAVINLVVVHSVFL
LTVPFRLTYLIKKTWMFGLPFCKFVSAMLHIHMYLTFLFYVVILVTRYLIFFKCKDKVEFYRKLHAVAASAGM
WTLVIVIVVPLVVSRYGIHEEYNBEHCFKFHKELAYTYVKIINYMIVIFVIAVAVILLVFQVFIIMLMVQKLR
HSLLSHQEFWAQLKNLFFIGVILVCFLPYQFFRIYYLNVVTHSNACNSKVAFYNEIFLSVTAISCYDLLLFVF
GGSHWFKQKIIGLWNCVLCR

The following DNA sequence nGPCR-40 <SEQ ID NO. 83> was identified in H. sapiens:

GCAGGAGCACTGAAAATCAGGAACAATCCTGTATTTTTTTGTGATAATCAACAAGGACAAAACTTCTCCATATG TAAATAACAGCGTTATGAGCAGCAATTCATCCCTGCTGGTGGCTGTGCAGCTGTGCTACGCGAACGTGAATGG GTCCTGTGTGAAAATCCCCTTCTCGCCGGGATCCCGGGTGATTCTGTACATAGTGTTTGGCTTTGGGGCTGTG CTGGCTGTTTTGGAAACCTCCTGGTGATGATTTCAATCCTCCATTTCAAGCAGCTGCACTCTCCGACCAATT GGTGGAGAGCTGCTGGTATTTTGGGAGGAGTTTTTGTACTTTCCACACCTGCTGTGATGTGGCATTTTGTTAC TCTTCTCTCTTTCACTTGTGCTTCATCTCCATCGACAGGTACATTGCGGTTACTGACCCCCTGGTCTATCCTA CCAAGTTCACCGTATCTGTGTCAGGAATTTGCATCAGCGTGTCCTGGATCCTGCCCCTCATGTACAGCGGTGC TGTGTTCTACACAGGTGTCTATGACGATGGGCTGGAGGAATTATCTGATGCCCTAAACTGTATAGGAGGTTGT TTCTGTATGGTAACATATTTCTTGTGGCTAGACGACAGGCGAAAAAGATAGAAAATACTGGTAGCAAGACAGA GTGGTAGCATTTATGATTTCATGGTTACCATATAGCATTGATTCATTAATTGATGCCTTTATGGGCTTTATAA CCCCTGCCTGTATTTATGAGATTTGCTGTTGGTGTGCTTATTATAACTCAGCCATGAATCCTTTGATTTATGC TTTATTTTACCCATGGTTTAGGAAAGCAATAAAAGTTATTGTAACTGGTCAGGTTTTAAAGAACAGTTCAGCA ACCATGAATTTGTTTTCTGAACATATATAA

The following amino acid sequence <SEQ ID NO. 84> is the predicted amino acid sequence derived from the DNA sequence of SEO ID NO. 83:

MSSNSSLLVAVQLCYANVNGSCVKIPFSPGSRVILYIVFGFGAVLAVFGNLLVMISILHFKQLHSPTNFLVAS
LACADFLVGVTVMPFSMVRTVESCWYFGRSFCTFHTCCDVAFCYSSLFHLCFISIDRYIAVTDPLVYPTKFTV

SVSGICISVSWILPLMYSGAVFYTGVYDDGLEELSDALNCIGGCQTVVNQNWVLTDFLSFFIPTFIMIILYGN
IFLVARRQAKKIENTGSKTESSSESYKARVARRERKAAKTLGVTVVAFMISWLPYSIDSLIDAFMGFITPACI
YEICCWCAYYNSAMNPLIYALFYPWFRKAIKVIVTGQVLKNSSATMNLFSEHI

The following DNA sequence nGPCR-54 <SEQ ID NO. 85> was identified in H. sapiens:

The following amino acid sequence <SEQ ID NO. 86> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 85:

MNEPLDYLANASDFPDYAAAFGNCTDENIPLKMHYLPVIYGIIFLVGFPGNAVVISTYIFKMRPWKSSTIIML NLACTDLLYLTSLPFLIHYYASGENWIFGDFMCKFIRFSFHFNLYSSILFLTCFSIFRYCVIIHPMSCFSIHK TRCAVVACAVVWIISLVAVIPMTFIITSTNRTNRSACLDLTSSDELNTIKWYNLILTASTFCLPLVIVTLCYT TIIHTLTHGLQTDSCLKQKARRLTILLLLAFYVCFLPFHILRVIQDRISACFQSVVPLRIRSMKLTSFLDHYA ALNTFGNLLYVVVSDNFOOAVCSTVRCKVSGNLEOAKKISYSN

The following DNA sequence nGPCR-56 <SEQ ID NO. 87> was identified in H. sapiens:

AAAAATTGCTGTACTGAACTATTGAATGGAACTTGGAAATAAAGTCCCTTCCAAAATAACTATTCTTCAACAG AGAGTAATAGGTAAATGTTTTAGAAGTGAGAGGACTCAAATTGCCAATGATTTACTCTTTTATTTTTCCTCCT AGGTTTCTGGGATAAGTATGTGCAAATAAAAATAAACATGAGAAGGAACTGTAACCTGATTATGGATTTGGG AAAAAGATAAATCAACACACAAAGGGAAAAGTAAACTGATTGACAGCCCTCAGGA**ATG**ATGCCCTTTTGCCAC AATATAATTAATATTTCCTGTGTGAAAAACAACTGGTCAAATGATGTCCGTGCTTCCCTGTACAGTTTAATGG TGCTCATAATTCTGACCACACTCGTTGGCAATCTGATAGTTATTGTTTCTATATCACACTTCAAACAACTTCA TACCCCAACAATTGGCTCATTCATTCCATGGCCACTGTGGACTTTCTTCTGGGGTGTCTGGTCATGCCTTAC AGTATGGTGAGATCTGCTGAGCACTGTTGGTATTTTGGAGAAGTCTTCTGTAAAATTCACACAAGCACCGACA ACTGAGATATAAAGCCAAGATGAATATCTTGGTTATTTGTGTGATGATCTTCATTAGTTGGAGTGTCCCTGCT GTTTTTGCATTTGGAATGATCTTTCTGGAGCTAAACTTCAAAGGCGCTGAAGAGATATATTACAAACATGTTC ACTGCAGAGGAGGTTGCTCTGTCTTTTAGCAAAATATCTGGGGTACTGACCTTTATGACTTCTTTTTATAT CCCTTTTCTTCACTACATTATTCCACCTACTTTGAATGATGTA

The following amino acid sequence <SEQ ID NO. 88> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 87:

MMPFCHNIINISCVKNNWSNDVRASLY<u>SLMVLIILTTLVGNLIVIVSISHF</u>KQLHTPTNWLIHSMATVDFLLG CLVMPYSMVRSAEHCWYFGEVFCKIHTSTDIMLSSASIFHLSFISIDRYYAVCDPLRYKAKMNILVICVMIFI

SWSVPAVFAFGMIFLELNFKGAEEIYYKHVHCRGGCSVFFSKISGVLTFMTSFYIPGSIMLCVYYRIYLIAKE QARLISDANQKLQIGLEMKNGISQSKERKAVKTLGIVMGVFLICWCPFFICTVMDPFLHYIIPPTLNDARGSR ANSA

The following DNA sequence nGPCR-56 <SEQ ID NO. 89> was identified in H. sapiens:

The following amino acid sequence <SEQ ID NO. 90> is the predicted amino acid sequence derived from the DNA sequence of SEO ID NO. 89:

MMPFCHNIINISCVKNNWSNDVRASLYSLMVLIILTTLVGNLIVIVSISHFKQLHTPTNWLIHSMATVDFLLG CLVMPYSMVRSAEHCWYFGEVFCKIHTSTDIMLSSASIFHLSFISIDRYYAVCDPLRYKAKMNILVICVMIFI SWSVPAVFAFGMIFLELNFKGAEEIYYKHVHCRGGCSVFFSKISGVLTFMTSFYIPGSIMLCVYYRIYLIAKE QARLISDANQKLQIGLEMKNGISQSKERKAVKTLGIVMGVFLICWCPFFICTVMDPFLHYIIPPTLNDVLIWF GYLNSTFNPMVYAFFYPWFRKALKMMLFGKIFQKDSSRCKLFLELSS

The following DNA sequence nGPCR-58 <SEQ ID NO. 91> was identified in H. sapiens:

CTGTAAAGTAGATTGTATGAGGACTCCATGAGGTCATCCACTTCAAGTCCTTGGCATAGGATAATTACTCAAA AGGTGATGACAATGGCGCAGGGAGGGATGGTGACTTGCCTGGAGATGCACAGCACCGTCTCTCCCATACTCGG TCATTCACACCATCATTGATTCACCAGGCACCACTCCGTGTCCAGCAGGACTCTGGGGACCCCAAATGGACAC TACCATGGAAGCTGACCTGGGTGCCACTGGCCACAGGCCCCGCACAGAGCTTGATGATGAGGACTCCTACCCC GGCTGGCCGGCTCCCAGGCCCGGCATGGAGCTGGCACGCGTCTGGCGCTGCTCCTGCTCAGCCTGGCCCTCTC TGACTTCTTGTTCCTGGCAGCAGCGGCCTTCCAGATCCTAGAGATCCGGCATGGGGGGACACTGGCCGCTGGGG ACAGCTGCCTGCCGCTTCTACTACTTCCTATGGGGCGTGTCCTACTCCTGCGGCCTCTTCCTGCTGGCCGCCC TCAGCCTCGACCGCTGCTGCCGCTGTGCCCACACTGGTACCCTGGGCACCGCCCAGTCCGCCTGCCCCT GTCTGGTGGTACGACCTGGTCATCTGCCTGGACTTCTGGGACAGCGAGGAGCTGTCGCTGAGGATGCTGGAGG TCCTGGGGGGCTTCCTGCCTTTCCTCCTGCTGCTCTGCCACGTGCTCACCCAGGCCACAGCCTGTCGCAC CTGCCACCGCCAACAGCAGCCCGCAGCCTGCCGGGGCCTTCGCCCGTGTGGCCAGGACCATTCTGTCAGCCTAT GTGGTCCTGAGGCTGCCCTACCAGCTGCCCAGCTGCTCTACCTGGGCCTTCCTGTGGGACGTCTACTCTGGCT ACCTGCTCTGGGAGGCCCTGGTCTACTCCGACTACTCTACTCAACAGCTGCCTCAGCCCCTTCCTCTG CCTCATGGCCAGTGCCGACCTCCGGACCCTGCTGCTCCTTCGCGGCAGCTCTCTGCGAG GAGCGGCCGGCCACTCACCCCACTGAGCCACAGACCCAGCTAGATTCTGAGGGTCCAACTCTGCCAGAGC CGATGGCAGAGGCCCAGTCACAGATGGATCCTGTGGCCCAGCCTCAGGTGAACCCCACACTCCAGCCACGATC GGATCCCACAGCTCAGCCACAGCTGAACCCTACGGCCCAGCCACAGTCGGATCCCACAGCCCAGCCACAGCTG AACCTCATGGCCCAGCCACAGTCAGATTCTGTGGCCCAGCCACAGGCAGACACTAACGTCCAGACCCCTGCAC CTGCTGCCAGTTCTGTGCCCAGTCCCTGTGATGAAGCTTCCCCAACCCCATCCTCGCATCCTACCCCAGGGGC CCTTGAGGACCCAGCCACACCTCCTGCCTCTGAAGGAGAAAGCCCCAGCAGCACCCCGCCAGAGGCGCCCCG GGAACCAGCCAGTCAGA

The following amino acid sequence <SEQ ID NO. 92> is the predicted amino

acid sequence derived from the DNA sequence of SEQ ID NO. 91:

LAWRCTAPSLPYSVIHTIIDSPGTTPCPAGLWGPOMDTTMEADLGATGHRPRTELDDEDSYPQGGWDTVFLVA LLLLGLPANGLMAWLAGSQARHGAGTRLALLLLSLALSDFLFLAAAAFQILEIRHGGHWPLGTAACRFYYFLW GVSYSSGLFLLAALSLDRCLLALCPHWYPGHRPVRLPLWVCAGVWVLATLFSVPWLVFPEAAVWWYDLVICLD FWDSEELSLMLEVLGGFLPFLLLLVCHVLTQATACRTCHRQQQPAACRGFARVARTILSAYVVLRLPYQLAQ LLYLAFLWDVYSGYLLWEALVYSDYLILLNSCLSPFLCLMASADLRTLLRSVLSSFAAALCEERPGSFTPTEP QTQLDSEGPTLPEPMAEAQSQMDPVAQPQVNPTLQPRSDPTAQPQLNPTAQPQSDPTAQPQLNLMAQPQSDSV AQPQADTNVQTPAPAASSVPSPCDEASPTPSSHPTPGALEDPATPPASEGESPSSTPPEAAPGAGPT

The following DNA sequence nGPCR-58 <SEQ ID NO. 93> was identified in H. sapiens:

ATGGACACTACCATGGAAGCTGACCTGGGTGCCACTGGCCACAGGGCCCCGCACAGAGCTTGATGATGAGGACT $\tt ATGGGTTGATGGCGTGGCCGGCTCCCAGGCCCGGCATGGAGCTGGCACGCGTCTGGCGCTGCTCCTGCT$ CAGCCTGGCCCTCTCTGACTTCTTGTTCCTGGCAGCAGCGGCCTTCCAGATCCTAGAGATCCGGCATGGGGGGA ${\tt CACTGGCCGCTGGGGACAGCTGCCTGCCGGCTTCTACTACTTCCTATGGGGCGTGTCCTACTCCTCCGGCCTCT}$ $\tt TTCCCCGAGGCTGCCGTCTGGTACGACCTGGTCATCTGCCTGGACTTCTGGGACAGCGAGGAGCTGTCGC$ TGAGGATGCTGGAGGTCCTGGGGGGGCTTCCTGCCTTTCCTCCTGCTGCTCGTCTCCCACGTGCTCACCCAGGC ATTCTGTCAGCCTATGTGGTCCTGAGGCTGCCCTACCAGCTGGCCAGCTGCTCTACCTGGCCTTCCTGTGGG ACGTCTACTCTGGCTACCTGCTCTGGGAGGCCCTGGTCTACTCCGACTACCTGATCCTACTCAACAGCTGCCT CAGCCCCTTCCTCTGCCTCATGGCCAGTGCCGACCTCCGGACCCTGCTGCTCCGTGCTCCTTCGCG GCAGCTCTCTGCGAGGGGCGGCCGGCCAGCTTCACGCCCACTGAGCCCAGACCCAGCTAGATTCTGAGGGTC CAACTCTGCCAGAGCCGATGGCAGAGGCCCAGTCACAGATGGATCCTGTGGCCCAGCCTCAGGTGAACCCCAC ACTCCAGCCACGATCGGATCCCACAGCTCAGCCACAGCTGAACCCTACGGCCCAGCCACAGTCGGATCCCACA GCCCAGCCACAGCTGAACCTCATGGCCCAGCCACAGTCAGACTCTGTGGCCCAGCCACAGGCAGACACTAACG TCCAGACCCCTGCACCTGCCAGTTCTGTGCCCAGTCCCTGTGATGAAGCTTCCCCAACCCCATCCTCGCA TCCTACCCCAGGGGCCCTTGAGGACCCAGCCACACCTCCTGCCTCTGAAGGAGAAAGCCCCAGCAGCACCCCG CCAGAGGCGCCCCGGGCGCAGGCCCCACGTGA

The following amino acid sequence <SEQ ID NO. 94 > 1s the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 93:

MDTTMEADLGATGHRPRTELDDEDSYPQGGWDTVFLVALLLLGLPANGLMAWLAGSQARHGAGTRLALLLLSL ALSDFLFLAAAAFQILEIRHGGHWPLGTAACRFYYFLWGVSYSSGLFLLAALSLDRCLLALCPHWYPGHRPVR LPLWVCAGVWVLATLFSVPWLVFPEAAVWWYDLVICLDFWDSEELSLRMLEVLGGFLPFLLLLVCHVLTQATA CRTCHRQQQPAACRGFARVARTILSAYVVLRLPYQLAQLLYLAFLWDVYSGYLLWEALVYSDYLILLNSCLSP FLCLMASADLRTLLRSVLSFFAAALCEERPGSFTPTEPQTQLDSEGPTLPEPMAEAQSQMDPVAQPQVNPTLQ PRSDPTAQPQLNPTAQPQSDYTAQPQLNLMAQPQSDSVAQPQADTNVQTPAPAA

The following DNA sequence nGPCR-3 <SEQ ID NO. 185> was identified in H sapiens:

The following amino acid sequence <SEQ ID NO. 186> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 185:

MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDMPFTLLGVMRGRTP SAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRPRYAGLLLGCAWGQSLAFSGAALGCSW LGYSSAFASCSLRLPPEPERPRFAAFTATLHAVGFVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLAD LHPSVRQRCLIQQKRRRHRATRKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPF TYSLLRRPFRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQQTH

EXAMPLE 2: CLONING OF nGPCR-X

To isolate a cDNA clone encoding full length nGPCR-x, a DNA fragment corresponding to a nucleotide sequence set forth in odd numbered nucleotide sequences ranging from SEQ ID NO: 1-93, or a portion thereof, can be used as a probe for hybridization screening of a phage cDNA library. The DNA fragment is amplified by the polymerase chain reaction (PCR) method. The PCR reaction mixture of 50 µl contains polymerase mixture (0.2 mM dNTPs, 1x PCR Buffer and 0.75 µl Expand High Fidelity Polymerase (Roche Biochemicals)), 1 µg of plasmid, and 50 pmoles of forward primer and 50 pmoles of reverse primer. The primers are preferably 10 to 25 nucleotides in length and are determined by procedures well known to those skilled in the art. Amplification is performed in an Applied Biosystems PE2400 thermocycler, using the following program: 95°C for 15 seconds, 52°C for 30 seconds and 72°C for 90 seconds; repeated for 25 cycles. The amplified product is separated from the plasmid by agarose gel electrophoresis, and purified by QiaquickTM gel extraction kit (Qiagen).

A lambda phage library containing cDNAs cloned into lambda ZAPII phage-vector is plated with *E. coli* XL-1 blue host, on 15 cm LB-agar plates at a density of 50,000 pfu per plate, and grown overnight at 37°C; (plated as described by Sambrook *et al.*, supra). Phage plaques are transferred to nylon membranes (Amersham Hybond NJ), denatured for 2 minutes in denaturation solution (0.5 M NaOH, 1.5 M NaCl), renatured for 5 minutes in renaturation solution (1 M Tris pH 7.5, 1.5 M NaCl), and washed briefly in 2xSSC (20x SSC: 3 M NaCl, 0.3 M Na-citrate). Filter membranes are dried and incubated at 80°C for 120 minutes to cross-link the phage DNA to the membranes.

The membranes are hybridized with a DNA probe prepared as described above. A DNA fragment (25 ng) is labeled with α -³²P-dCTP (NEN) using

5

10

20

RediprimeTM random priming (Amersham Pharmacia Biotech), according to manufacturers instructions. Labeled DNA is separated from unincorporated nucleotides by S200 spin columns (Amersham Pharmacia Biotech), denatured at 95°C for 5 minutes and kept on ice. The DNA-containing membranes (above) are prehybridized in 50 ml ExpressHybTM (Clontech) solution at 68°C for 90 minutes. Subsequently, the labeled DNA probe is added to the hybridization solution, and the probe is left to hybridize to the membranes at 68°C for 70 minutes. The membranes are washed five times in 2x SSC, 0.1% SDS at 42°C for 5 minutes each, and finally washed 30 minutes in 0.1x SSC, 0.2% SDS. Filters are exposed to Kodak XARTM film (Eastman Kodak Company, Rochester, N.Y., USA) with an intensifying screen at -80°C for 16 hours. One positive colony is isolated from the plates, and replated with about 1000 pfu on a 15 cm LB plate. Plating, plaque lift to filters and hybridization are performed as described above. About four positive phage plaques are isolated form this secondary screening.

cDNA containing plasmids (pBluescript SK-) are rescued from the isolated phages by in vivo excision by culturing XL-1 blue cells co-infected with the isolated phages and with the Excision helper phage, as described by manufacturer (Stratagene). XL-blue cells containing the plasmids are plated on LB plates and grown at 37°C for 16 hours. Colonies (18) from each plate are replated on LB plates and grown. One colony from each plate is stricken onto a nylon filter in an ordered array, and the filter is placed on a LB plate to raise the colonies. The filter is then hybridized with a labeled probe as described above. About three positive colonies are selected and grown up in LB medium. Plasmid DNA is isolated from the three clones by Qiagen Midi KitTM (Qiagen) according to the manufacturer's instructions. The size of the insert is determined by digesting the plasmid with the restriction enzymes *Notl* and *Sall*, which establishes an insert size. The sequence of the entire insert is determined by automated sequencing on both strands of the plasmids.

nGPCR-1: PCR AND SUBCLONING

cDNAs were sequenced directly using an AB1377 fluorescence-based sequencer (Perkin Elmer/Applied Biosystems Division, PE/ABD, Foster City, CA) and the ABI PRISM Ready Dye-Deoxy Terminator kit with Taq FS polymerase. Each ABI cycle sequencing reaction contained about 0.5µg of plasmid DNA. Cycle-sequencing was performed using an initial denaturation at 98°C for 1 min, followed

5

10

15

20

25

by 50 cycles: 98°C for 30 sec, annealing at 50°C for 30 sec, and extension at 60°C for 4 min. Temperature cycles and times were controlled by a Perkin-Elmer 9600 thermocycler. Extension products were purified using AGTC® gel filtration block (Edge BiosSystems, Gaithersburg, MD). Each reaction product was loaded by pipette onto the column, which was then centrifuged in a swinging bucket centrifuge (Sorvall model RT6000B table top centrifuge) at 1500 x g for 4 min at room temperature. Column-purified samples were dried under vacuum for about 40 min and then dissolved in 5µl of a DNA loading solution (83% deionized formamide, 8.3 mM EDTA, and 1.6 mg/ml Blue Dextran). The samples were then heated to 90°C for three min and loaded into the gel sample wells for sequence analysis by the ABI377 sequencer. Sequence analysis was performed by importing ABI373A files into the Sequencher program (Gene Codes, Ann Arbor, MI).

The PCR reaction was performed in 50μ L samples containing 41.9μ L H₂O, 5μ L 10x Buffer containing 15 mM MgCl₂ (Boehringer Mannheim Expand High Fidelity PCR System), 0.5μ L 10mM dNTP mix, 1.5μ L human genomic DNA (Clontech #6550-1, 0.1μ g/ μ L), 0.3μ L primer VR1A (1μ g/ μ L), 0.3μ L primer VR1B (1μ g/ μ L), and 0.5μ L High Fidelity Taq polymerase (Bochringer Mannheim, 3.5U/ μ l). The primer sequences for and, respectively were: 5TCAAAGCTTATGGAATCATCTTTCTCATTTGGAGTGATCCTTGCTGTC,

(VR1A)(SEQ ID NO: 95) corresponding to the 5' end of the coding region and containing a *HindIII* restriction site, and:

5' TTCACTCGAGTTAGCCATCAAACTCTGAGCTGGAGATAGTGACGATGTG (VR1B)(SEQ ID NO: 96) corresponding to the 3' end of the coding region and containing an *Xho*I restriction site (Genosys). The PCR reaction was carried out using a GeneAmp PCR9700 thermocycler (Perkin Elmer Applied Biosystems) and started with 1 cycle of 94°C for 2 min followed by 5 cycles at 94°C for 30 sec, 60°C for 2 min, 72°C for 1.5 min, followed by 20 cycles at 94°C for 30 sec, 60°C for 30 sec, 72°C for 1.5 min.

The PCR reaction was loaded onto a 0.75% agarose gel. The DNA band was excised from the gel and the DNA eluted from the agarose using a QIAquick gel extraction kit (Qiagen). The eluted DNA was ethanol-precipitated and resuspended in 4μ L H₂O for ligation. The ligation reaction consisted of 4μ L of fresh ethanol-precipitated PCR product and 1μ L of pCRII-TOPO vector (Invitrogen). The reaction was gently mixed and allowed to incubate for 5 min. at room temperature followed by

5

10

15

20

25

the addition of $1\mu L$ of 6x TOPO cloning stop solution and mixing for 10 sec. at room temperature. The sample was then placed on ice and $2\mu L$ was transformed in $50\mu L$ of One Shot cells (Invitrogen) and plated onto ampicillin plates. Four white colonies were chosen and the presence of an insert was verified by PCR in the following manner. Each colony was resuspended in 2 ml LB broth for 2 hrs. A $500\mu L$ aliquot was spun down in the microfuge, the supernatant discarded, and the pellet resuspended in $25\mu L$ of H_2O . A $16\mu L$ aliquot was removed and boiled for 5 min and the sample was placed on ice. The sample was microfuged briefly to pellet any bacterial debris and PCR was carried out with $15\mu L$ sample using primers VR1A and VR1B, described above.

Colonies from positive clones identified by PCR were used to inoculate a 4ml culture of LB medium containing 100 μ g/ml ampicillin. Plasmid DNA was purified using the Wizard Plus Minipreps DNA purification system (Promega). Since the primers used to amplify the fragment of nGPCR-1 from genomic DNA were engineered to have *HindIII* and *XhoI* sites, the cDNA obtained from the minipreps was digested with these restriction enzymes. One clone was verified by gel electrophoresis to give a DNA band of the correct size. cDNA from this clone was then sequenced, yielding the sequence of SEQ ID NO: 73.

nGPCR-3: PCR AND SUBCLONING

First-strand cDNA synthesis was performed following the directions for 3'-RACE ready cDNA from the SMARTTM RACE cDNA Amplification Kit (Clontech). First 3 μ l of H₂O, 1 μ l human whole brain poly A⁺ RNA (1 μ g/ μ l) (Clontech, 6516-1) and 1 μ l 3'-CDS primer were mixed together, incubated at 70°C for 2 minutes, then placed on ice for 2 minutes. Added to the tube was 2 μ l 5X First-Strand buffer, 1 μ l 20 mM DTT, 1 μ l dNTP mix (10 mM) and 1 μ l Superscript II RT (200 units/ μ l) (GIBCO/BRL). The tube was incubated at 42°C for 1.5 hours then the reaction was diluted with 250 μ l of Tricine-EDTA buffer.

PCR was performed in a 50 μl reaction using components that come with the Advantage®-GC cDNA PCR Kit. The PCR reaction contained 22.4 μl H₂O, 10 μl 5X GC cDNA PCR Reaction buffer, 10 μl 5M GC Melt, 1μl 50X dNTP mix (10 mM each), 5 μl human brain cDNA, 0.3 μl of LW1649 (SEQ ID NO: 187)(1 μg/μl), 0.3 μl of LW1650 (SEQ ID NO: 188)(1 μg/μl), 1 μl 50X Advantage-GC cDNA polymerase mix. The PCR reaction was performed in a Perkin-Elmer 9600 GeneAmp PCR

5

10

15

20

25

System starting with 1 cycle of 94°C for 2 min then 8 cycles at 94°C for 15 sec, 72°C for 2 min (decreasing 1°C with each cycle), 72°C for 3 min, followed by 30 cycles of 94°C for 15 sec, 68°C for 3 min. The PCR reaction was loaded onto a 1.2 % agarose gel. The DNA band was excised from the gel, placed in GenElute Agarose spin column (Supelco) and spun for 10 min at maximum speed in a microcentrifuge. The eluted DNA was EtOH precipitated and resuspended in 4 H₂O for ligation. The PCR primer sequence for LW1649 was:

GCATAAGCTTGCCATGGGCCCCGGCGAGG (SEQ ID NO: 187) and for LW1650 was:

GCATTCTAGA<u>CCTCAGTGTGTCTGCTGC</u> (SEQ ID NO: 188). The underlined portion of the primers matches the 5' and 3' areas, respectively, of the coding region.

The ligation reaction used solutions from the TOPO TA Cloning Kit (Invitrogen) which consisted of 4μ l PCR product DNA, 1μ l Salt Solution and 1μ l pCRII-TOPO vector that was incubated for 5 minutes at room temperature and then placed on ice. Two microliters of the ligation reaction was transformed in One-Shot TOP10 cells (Invitrogen), and placed on ice for 30 minutes. The cells were heat-shocked for 30 seconds at 42° C, placed on ice for two minutes, 250μ l of SOC was added, then incubated at 37° C with shaking for one hour and then plated onto ampicillin plates. A single colony containing an insert was used to inoculate a 5 ml culture of LB medium. Plasmid DNA was purified using a Concert Rapid Plasmid Miniprep System (GibcoBRL) and then sequenced.

The DNA subcloned into pCRII-TOPO was sequenced using the ABI PRISMTM 310 Genetic Analyzer (PE Applied Biosystems) which uses advanced capillary electrophoresis technology and the ABI PRISMTM BigDyeTM Terminator Cycle Sequencing Ready Reaction Kit. Each cycle-sequencing reaction contained 6 µl of H₂0, 8 µl of BigDye Terminator mix, 5 µl mini-prep DNA (0.1 µg/µl), and 1 µl primer (25 ng/µl) and was performed in a Perkin-Elmer 9600 thermocycler with 25 cycles of 96°C for 10 sec, 50°C for 10 sec, and 60°C for 4 min. The product was purified using a CentriflexTM gel filtration cartridge, dried under vacuum, then dissolved in 16 µl of Template Suppression Reagent (PE Applied Biosystems). The samples were heated at 95°C for 5 min then placed in the 310 Genetic Analyzer, yielding the sequence of SEQ ID NO: 95.

5

10

15

20

25

nGPCR-9: PCR AND SUBCLONING

The PCR reaction was performed in 50 μ l containing 34.5 μ l H₂O, 5 μ l Buffer II (PE Applied Biosystems AmpliTaq Gold system), 6 μ l 25 mM MgCl₂, 2 μ l 10 mM dNTP mix, 1.5 μ l human genomic DNA (Clontech #6550-1, 0.1 μ g/ μ l), 0.3 μ l primer VR9A (1 μ g/ μ l), 0.3 μ l primer VR9B (1 μ g/ μ l), and 0.4 μ l AmpliTaq GoldTM DNA Polymerase. The primer sequences for VR9A and VR9B were as follows:

VR9A 5'TTCAAAGCTTATGGAGTCGGGGCTGCTG 3' (SEQ ID NO: 101), corresponding to the 5' end of the coding region and containing a *HindIII* restriction site, and the reverse primer was:

VR9B 5' TTCACTCGAGTCAGTCTGCAGCCGGTTCTG 3', (SEQ ID NO: 102), corresponding to the 3' end of the coding region and containing an XhoI restriction site (Genosys). The PCR reaction was carried out using a GeneAmp PCR 9700 thermocycler (Perkin Elmer Applied Biosystems) and started with 1 cycle of 95°C for 10 min, then 10 cycles at 95°C for 30 sec, 72°C for 2 min decreasing 1°C each cycle, 72°C for 1 min, followed by 30 cycles at 95°C for 30 sec, 60°C for 30 sec, 72°C for 1 min. The PCR reaction was loaded on a 0.75% gel. The DNA band was excised from the gel and the DNA was eluted from the agarose using a QIAquick gel extraction kit (Qiagen). The eluted DNA was ethanol-precipitated and resuspended in 4 μl H₂O for ligation. The ligation reaction consisted of 4 μl of fresh ethanolprecipitated PCR product and 1 µl of pCRII-TOPO vector (Invitrogen). The reaction was gently mixed and allowed to incubate for 5 min at room temperature followed by the addition of 1 µl of 6x TOPO cloning stop solution and mixing for 10 sec at room temperature. The sample was then placed on ice and 2 µl was transformed in 50 µl of One Shot cells (Invitrogen) and plated onto ampicillin plates. Five white colonies were chosen and were used to inoculate a 4 ml culture of LB medium containing 100 ug/ml ampicillin. Plasmid DNA was purified using the Wizard Plus Minipreps DNA purification system (Promega). Since the primers used to PCR SEQ-9 from genomic DNA were engineered to have HindIII and XhoI sites, the cDNA obtained from the minipreps was digested with these restriction enzymes. One clone was verified by gel electrophoresis to give a DNA band of the correct size. cDNA from this clone was then submitted for sequencing. One mutation was found (bp 621 T→G) and repaired as described as below.

5

10

15

20

25

The mutation in the identified clone was repaired using the QuikChange Site-Directed Mutagenesis Kit (Stratagene). The PCR reaction contained 39.3 μl H₂O, 5 μl 10x reaction buffer, 50 ng mini-prep cDNA, 1.25 μl primer VR9E (100 ng/μl), 1.25 μl primer VR9F (100 ng/μl), 1 μl 20 mM dNTP mix, 1 μl Pfu DNA polymerase. The cycle conditions were 95°C for 30 sec, then 12 cycles at 95°C for 30 sec, 55°C for 1 min, 68°C for 10 min. One μl of DpnI was added and the tube incubated at 37°C for 1 hr. One μl of the DpnI-treated DNA was transformed into 50 μl Epicurian coli XL1-Blue supercompetent cells and the entire insert was re-sequenced. The primer sequences used were:

VR9E: 5' GCATCCTGGCCGC $\underline{\mathbf{T}}$ ATCTGTGCACTCTACG 3' (SEQ ID NO: 103) and

VR9F: 5' CGTAGAGTGCACAGATAGCGGCCAGGATGC 3' (SEQ ID NO: 104) where the base underlined was the base being corrected.

The clone described above was sequenced directly using an ABI377 fluorescence-based sequencer (Perkin Elmer/Applied Biosystems Division, PE/ABD, Foster City, CA) and the ABI BigDyeTM Terminator Cycle Sequencing Ready Reaction kit with Taq FSTM polymerase. Each ABI cycle sequencing reaction contained 0.5 µg of plasmid DNA. Cycle-sequencing was performed using an initial denaturation at 98°C for 1 min, followed by 50 cycles: 96°C for 30 sec, annealing at 50°C for 30 sec, and extension at 60°C for 4 min. Temperature cycles and times were controlled by a Perkin-Elmer 9600 thermocycler. Extension products were purified using AGTC (R) gel filtration block (Edge BiosSystems, Gaithersburg, MD). Each reaction product was loaded by pipette onto the column, which was then centrifuged in a swinging bucket centrifuge (Sorvall model RT6000B tabletop centrifuge) at 1500 x g for 4 min at room temperature. Column-purified samples were dried under vacuum for about 40 min and then dissolved in 3 µl of a DNA loading solution (83% deionized formamide, 8.3 mM EDTA, and 1.6 mg/ml Blue Dextran). The samples were then heated to 90°C for 3.5 min and loaded into the gel sample wells for sequence analysis by the ABI377 sequencer. Sequence analysis was performed by importing ABI377 files into the 310 Genetic Analyzer, yielding the sequence of SEQ ID NO: 77.

5

10

15

20

25

nGPCR-11: PCR AND SUBCLONING

PCR was performed in a 50 μ l reaction containing 32 μ l H₂O, 5 μ l 10X TT buffer (140 mM Ammonium Sulfate, 0.1 % gelatin, 0.6 M Tris-tricine pH 8.4), 5 μ l 15 mM MgS04, 2 μ l 10 mM dNTP, 5 μ l human genomic DNA (0.3 μ g/ μ l)(Clontech), 0.3 μ l of LW1564 (1 μ g/ μ l), 0.3 μ l of LW1565 (1 μ g/ μ l), 0.4 μ l High Fidelity Taq polymerase (Boehringer Mannheim). The PCR reaction was performed in a GeneAmp 9600 PCR thermocycler (PE Applied Biosystems) starting with 1 cycle of 94°C for 2 min followed by 17 cycles at 94°C for 30 sec, 72°C for 2 min decreasing 1°C each cycle, 68°C for 2 min, then 25 cycles of 94°C for 30 sec, 55°C for 30 sec, 68°C for 2 min. The PCR reaction was loaded onto a 1.2 % agarose gel. The DNA band was excised from the gel, placed in GenElute Agarose spin column (Supelco) and spun for 10 min at maximum speed in a microcentrifuge. The eluted DNA was EtOH precipitated and resuspended in 4 μ l H₂O for ligation. The forward PCR primer sequence was:

LW1564: GCATAAGCTTCCATGTACAACGGGTCGTGCTGC (SEQ ID NO: 107), and the reverse PCR primer was:

 $LW1565: GCATTCTAGATCAGTGCCACTCAACAATGTGGG \ (SEQ\ ID\ NO:\ 108).$

The ligation reaction used solutions from the TOPO TA Cloning Kit (Invitrogen) which consisted of 4 µl PCR product DNA and 1 µl pCRII-TOPO vector that was incubated for 5 minutes at room temperature. To the ligation reaction one microliter of 6X TOPO Cloning Stop Solution was added then the reaction was placed on ice. Two microliters of the ligation reaction was transformed in One Shot TOP10 cells (Invitrogen), and placed on ice for 30 minutes. The cells were heat-shocked for 30 seconds at 42°C, placed on ice for two minutes, 250 Tl of SOC was added, then incubated at 37°C with shaking for one hour and then plated onto ampicillin plates. A single colony containing an insert was used to inoculate a 5 ml culture of LB medium. Plasmid DNA was purified using a Concert Rapid Plasmid Miniprep System (GibcoBRL) and then sequenced.

The DNA subcloned into pCRII was sequenced using the ABI PRISMTM 310 Genetic Analyzer (PE Applied Biosystems) which uses advanced capillary electrophoresis technology and the ABI PRISMTM BigDyeTM Terminator Cycle Sequencing Ready Reaction Kit. Each cycle-sequencing reaction contained 6 µl of

5

10

15

20

25

 H_20 , 8 μ l of BigDye Terminator mix, 5 μ l mini-prep DNA (0.1 μ g/ μ l), and 1 μ l primer (25 ng/ μ l) and was performed in a Perkin-Elmer 9600 thermocycler with 25 cycles of 96°C for 10 sec, 50°C for 10 sec, and 60°C for 4 min. The product was purified using a CentriflexTM gel filtration cartridge, dried under vacuum, then dissolved in 16 μ l of Template Suppression Reagent (PE Applied Biosystems). The samples were heated at 95°C for 5 min then placed in the 310 Genetic Analyzer, yielding the sequence of SEQ ID NO: 79.

nGPCR-16: PCR AND SUBCLONING

PCR was performed in a 50 μl reaction containing 32 μl H₂O, 5 μl 10X TT buffer (140 mM Ammonium Sulfate, 0.1 % gelatin, 0.6 M Tris-tricine pH 8.4), 5 μl 15 mM MgS0₄, 2 μl 10 mM dNTP, 5 μl 2445704H1 DNA (0.17 Tg/Tl), 0.3 μl of LW1587 (1 μg/μl), 0.3 μl of LW1588 (1 μg/μl), 0.4 μl High Fidelity Taq polymerase (Boehringer Mannheim). The PCR reaction was performed on a Robocycler thermocycler (Stratagene) starting with 1 cycle of 94°C for 2 min followed by 15 cycles of 94°C for 30 sec, 55°C for 1.3 min, 68°C for 2 min. The PCR reaction was loaded onto a 1.2 % agarose gel. The DNA band was excised from the gel, placed in GenElute Agarose spin column (Supelco) and spun for 10 min at maximum speed in a microcentrifuge. The eluted DNA was EtOH precipitated and resuspended in 12μl H₂O for ligation. The PCR primer sequence for the forward primer was:

LW1587: GATCAAGCTTATGACAGGTGACTTCCCAAGTATGC (SEQ ID NO: 111), and the sequence for the reverse primer was:

LW1588: GATCCTCGAGGCTAACGGCACAAAACACAATTCC (SEQ ID NO: 112).

The ligation reaction used solutions from the TOPO TA Cloning Kit (Invitrogen) which consisted of 4µl PCR product DNA and 1 µl pCRII-TOPO vector that was incubated for 5 minutes at room temperature. To the ligation reaction one microliter of 6X TOPO Cloning Stop Solution was added then the reaction was placed on ice. Two microliters of the ligation reaction was transformed in One-Shot TOP10 cells (Invitrogen), and placed on ice for 30 minutes. The cells were heat-shocked for 30 seconds at 42°C, placed on ice for two minutes, 250 µl of SOC was added, then incubated at 37°C with shaking for one hour and then plated onto ampicillin plates. A single colony containing an insert was used to inoculate a 5 ml culture of LB medium.

5

10

15

20

25

Plasmid DNA was purified using a Concert Rapid Plasmid Miniprep System (GibcoBRL) and then sequenced.

The DNA subcloned into pCRII was sequenced using the ABI PRISMTM 310 Genetic Analyzer (PE Applied Biosystems) which uses advanced capillary electrophoresis technology and the ABI PRISMTM BigDyeTM Terminator Cycle Sequencing Ready Reaction Kit. Each cycle-sequencing reaction contained 6 μl of H₂0, 8 μl of BigDye Terminator mix, 5 μl mini-prep DNA (0.1 μg/μl), and 1 μl primer (25 ng/μl) and was performed in a Perkin-Elmer 9600 thermocycler with 25 cycles of 96°C for 10 sec, 50°C for 10 sec, and 60°C for 4 min. The product was purified using a CentriflexTM gel filtration cartridge, dried under vacuum, then dissolved in 16 μl of Template Suppression Reagent (PE Applied Biosystems). The samples were heated at 95°C for 5 min then placed in the 310 Genetic Analyzer, yielding the sequence of SEQ ID NO: 81.

nGPCR-40: PCR AND SUBCLONING

PCR was performed in a 50 µl reaction containing utilizing Herculase DNA Polymerase blend (Stratagene), using the buffer recommendations provided by the manufacturer, 200 ng each of primers PSK 18 and 19 (SEQ ID NOS: 115 and 116), 150 ng of human genomic DNA (Clontech), and 2% DMSO. The PCR reaction was performed on a Robocycler thermocycler (Stratagene) starting with 1 cycle of 94°C for 2 min followed by 35 cycles of 94°C for 30 sec, 65°C for 30 sec, 72°C for 2 min. The PCR reaction was purified using the QiaQuick PCR Purification Kit (Qiagen), and then eluted in TE. The PCR primer sequences were:

PSK 18 GATC GAATTCGCAGGAGCAATG AAAATCAGGAAC (SEQ ID NO: 115), and:

PSK19: GATCGAATTC<u>TTATATATGTTCAGAAAACAAATTCATGG</u> (SEQ ID NO: 116)). The underlined portion of the primer matches the 5' and 3' areas, respectively, of a portion of the 5' untranslated region and coding region. Initiation and termination codons are shown above in bold.

The PCR product was ligated into the pCR-BluntII-TOPO vector (Invitrogen) using the Zero Blunt Topo PCR TA cloning kit as follow: $3\mu l$ PCR product DNA, $1\mu l$ pCRII-TOPO vector, and $1\mu l$ TOPOII salt solution (1.2M NaCl, 0.06M MgCl₂). The mixture was incubated for 5 minutes at room temperature. To the ligation reaction one microliter of 6X TOPO Cloning Stop Solution was added, and then the

5

10

15

20

25

reaction was placed on ice. Two microliters of the ligation reaction was transformed in One-Shot TOP10 cells (Invitrogen), and placed on ice for 30 minutes. The cells were heat-shocked for 30 seconds at 42° C, placed on ice for two minutes, 250μ l of SOC was added, then incubated at 37° C with shaking for one hour and then plated onto ampicillin plates supplemented with Xgal and IPTG. Single colonies were screened by PCR for the presence of the insert, and a plasmid DNA from colony 58 was purified using a Qiagen Endo-Free plasmid purification kit.

5

10

15

20

25

30

NSDOCID KWO 0136473A2 L >

nGPCR-40 was sequenced directly using an ABI377 fluorescence-based sequencer (Perkin Elmer/Applied Biosystems Division, PE/ABD, Foster City, CA) and the ABI BigDveTM Terminator Cycle Sequencing Ready Reaction kit with Tag FSTM polymerase. Each ABI cycle sequencing reaction contained about 0.5 ug of plasmid DNA. Cycle-sequencing was performed using an initial denaturation at 98°C for 1 min, followed by 50 cycles: 96°C for 30 sec, annealing at 50°C for 30 sec, and extension at 60°C for 4 min. Temperature cycles and times were controlled by a Perkin-Elmer 9600 thermocycler. Extension products were purified using AGTC® gel filtration block (Edge BiosSystems, Gaithersburg, MD). Each reaction product was loaded by pipette onto the column, which was then centrifuged in a swinging bucket centrifuge (Sorvall model RT6000B tabletop centrifuge) at 1500 x g for 4 min at room temperature. Column-purified samples were dried under vacuum for about 40 min and then dissolved in 3 μl of DNA loading solution (83% deionized formamide, 8.3 mM EDTA, and 1.6 mg/ml Blue Dextran). The samples were then heated to 90°C for 3.5 min and loaded into the gel sample wells for sequence analysis by the ABI377 sequencer. Sequence analysis was performed by importing ABI377 files into the Sequencher program (Gene Codes, Ann Arbor, MI), which yielded a sequence identical to SEQ ID NO:83 with the exception that the nucleotide at position 10 was identified as an "A" which incorrectly indicated the presence of an initiation codon at that position. Subsequent analysis of genomic DNA samples indicated that this position was incorrectly assigned and that the correct nucleotide at that position was a "C". The sequence reported at SEQ ID NO. 83 correctly identifies the nucleotide at position 10 and indicates that the first initiation codon occurs at position 88-90.

nGPCR-54: PCR AND SUBCLONING

Two microliters of a human genomic library (~10⁸ PFU/ml) (Clontech) was added to 6 ml of an overnight culture of K802 cells (Clontech), then distributed as 250 µl aliquots into each of 24 tubes. The tubes were incubated at 37°C for 15 min. Seven milliliters of 0.8% agarose was added to each tube, mixed, then poured onto LB agar + 10 mM MgSO₄ plates and incubated overnight at 37°C. To each plate 5 ml of SM (0.1M NaCl, 8.1 mM MgSO₄-7H₂O, 50mM Tris-Cl (pH 7.5), 0.0001% gelatin) phage buffer was added and the top agarose was removed with a microscope slide and placed in a 50 ml centrifuge tube. A drop of chloroform was added and the tube was place in a 37 °C shaker for 15 min, then centrifuged for 20 min at 4000 RPM (Sorvall RT6000 table top centrifuge) and the supernatant stored at 4°C as a stock solution.

Two ul of phage from each tube was heated to 99°C for 4 min then cooled to 10°C. Added to the phage was a PCR mix containing 8.8 µl H₂O, 4 µl 5X Rapid-Load Buffer (Origene), 2 µl 10xPCR buffer II (Perkin-Elmer), 2 µl 25 mM MgCl₂, 0.8 µl 10 mM dNTP, 0.12 µl LW1634 (1 µg/µl)(SEQ ID NO: 119), 0.12 µl LW1635 (1 μg/μl)(SEQ ID NO: 120), 0.2 μl AmpliTaq Gold polymerase (Perkin Elmer). The PCR reaction involved 1 cycle at 95°C for 10 min followed by 35 cycles at 95°C for 45 sec, 53.5°C for 2 min, 72°C for 45 sec. The reaction was loaded onto a 2 % agarose gel. From the tube that gave a PCR product of the correct size, 10 µl was used to set up five 1:10 dilutions that were plated onto LB agar + 10 mM MgSO₄ plates and incubated overnight. A BA85 nitrocellulose filter (Schleicher & Schuell) was placed on top of each plate for 1 hour. The filter was removed, placed phage side up in a petri dish, and covered with 4 ml of SM for 15 min to elute the phage. One milliliter of SM was removed from each plate and used to set up a PCR reaction as above. The plate of the lowest dilution to give a PCR product was subdivided, filterlifted and the PCR reaction was repeated. The series of dilutions and subdividing of the plate was continued until a single plaque was isolated that gave a positive PCR band. Once a single plaque was isolated, 10 µl phage supernatant was added to 100 ul SM and 200 ul of K802 cells per plate with a total of 8 plates set up. The plates were incubated overnight at 37°C. The top agarose was removed by adding 8 ml of SM then scrapping off the agarose with a microscope slide and collected in a centrifuge tube. To the tube, 3 drops of chloroform was added, vortexed, incubated at 37°C for 15 min then centrifuged for 20 min at 4000 RPM (Sorvall RT6000 table top

10

15

20

25

centrifuge) to recover the phage, which was used to isolate genomic phage DNA using the Qiagen Lambda Midi Kit. The sequence for primer LW1634 was:

CTGAAAGTTGTCGCTGACC (SEQ ID NO: 119), and for primer LW1635 was:

CGATTATCCACACTTTGACCC (SEQ ID NO: 120).

The PCR reaction for the coding region was performed in a 50 μ l reaction containing 33 μ l H₂O, 5 μ l 10X TT buffer (140 mM Ammonium Sulfate, 0.1 % gelatin, 0.6 M Tris-tricine pH 8.4), 5 μ l 15 mM MgS0₄, 2 μ l 10 mM dNTP, 4 μ l genomic phage DNA (0.25 μ g/ μ l), 0.3 μ l LW1698 (1 μ g/ μ l)(SEQ ID NO: 121), 0.3 μ l LW1699 (1 μ g/ μ l)(SEQ ID NO: 122), 0.4 μ l High Fidelity Taq polymerase (Boehringer Mannheim). The PCR reaction was started with 1 cycle of 94°C for 2 min followed by 30 cycles at 94°C for 30 sec., 55°C for 30 sec., 68°C for 2 min. The PCR reaction was loaded onto a 2 % agarose gel. The DNA band was excised from the gel, placed in GenElute Agarose spin column (Supelco) and spun for 10 min at maximum speed. The eluted DNA was EtOH precipitated and resuspended in 8 μ l H₂O. The PCR primer sequence for primer LW1698 was:

GCATACCATGAATGAGCCACTAGAC (SEQ ID NO: 121), and for primer LW1699 was:

GCATCTCGAGTCAAGGGTTGTTTGAGTAAC (SEQ ID NO: 122). The underlined portion of the primer matches the 5' and 3' areas, respectively, of the coding region of nGPCR-54.

The ligation reaction used solutions from the TOPO TA Cloning Kit (Invitrogen) which consisted of 4µl PCR product DNA, 1 µl of salt solution and 1 µl pCRII-TOPO vector that was incubated for 5 minutes at room temperature then the reaction was placed on ice. Two microliters of the ligation reaction was transformed in One-Shot TOP10 cells (Invitrogen), and placed on ice for 30 minutes. The cells were heat-shocked for 30 seconds at 42°C, placed on ice for two minutes, 250 µl of SOC was added, then incubated at 37°C with shaking for one hour and then plated onto ampicillin plates. A single colony containing an insert was used to inoculate a 5 ml culture of LB medium. Plasmid DNA was purified using a Concert Rapid Plasmid Miniprep System (GibcoBRL) and then sequenced.

nGPCR-54 genomic phage DNA was sequenced using the ABI PRISMTM 310 Genetic Analyzer (PE Applied Biosystems) which uses advanced capillary

5

10

15

20

25

electrophoresis technology and the ABI PRISMTM BigDyeTM Terminator Cycle Sequencing Ready Reaction Kit. The cycle-sequencing reaction contained 14 μ l of H₂0, 16 μ l of BigDye Terminator mix, 7 μ l genomic phage DNA (0.1 μ g/ μ l), and 3 μ l primer (25 ng/ μ l). The reaction was performed in a Perkin-Elmer 9600 thermocycler at 95°C for 5 min, followed by 99 cycles of 95°C for 30 sec, 55°C for 20 sec, and 60°C for 4 min. The product was purified using a CentriflexTM gel filtration cartridges, dried under vacuum, then dissolved in 16 μ l of Template Suppression Reagent. The samples were heated at 95°C for 5 min then placed in the 310 Genetic Analyzer.

The DNA subcloned into pCRII was sequenced using the ABI PRISMTM 310 Genetic Analyzer (PE Applied Biosystems) which uses advanced capillary electrophoresis technology and the ABI PRISMTM BigDyeTM Terminator Cycle Sequencing Ready Reaction Kit. Each cycle-sequencing reaction contained 6 μl of H₂0, 8 μl of BigDye Terminator mix, 5 μl mini-prep DNA (0.1 μg/μl), and 1 μl primer (25 ng/μl) and was performed in a Perkin-Elmer 9600 thermocycler with 25 cycles of 96°C for 10 sec, 50°C for 10 sec, and 60°C for 4 min. The product was purified using a CentriflexTM gel filtration cartridge, dried under vacuum, then dissolved in 16 μl of Template Suppression Reagent (PE Applied Biosystems). The samples were heated at 95°C for 5 min then placed in the 310 Genetic Analyzer, yielding the sequence of SEQ ID NO: 85.

nGPCR-56: PCR AND SUBCLONING

The PCR reaction for the coding region of nGPCR-56 used components that come with PLATINUM® *Pfx* DNA Polymerase (GibcoBRL) containing 35.5 μl H₂O, 5 μl 10X Pfx Amplification buffer, 1.5 μl 50mM MgSO₄, 2 μl 10 mM dNTP, 5 μl human genomic DNA (0.3μg/μl)(Clontech), 0.3 μl of LW1603 (1 μg/μl)(SEQ ID NO: 152), 0.3 μl of LW1604 (1 μg/μl)(SEQ ID NO: 153), 0.4 μl PLATINUM® *Pfx* DNA Polymerase (2.5 U/Tl). The PCR reaction was performed in a Robocycler Gradient 96 (Stratagene) starting with 1 cycle of 94°C for 5 min followed by 30 cycles at 94°C for 40 sec, 55°C for 2 min, 68°C for 3 min. Following the final cycle, 0.5 μl of AmpliTaq DNA Polymerase (5 U/μl) was added and the tube was incubated at 72°C for 5 min. The sequence of LW1603 is:

5

10

15

20

25

GATCAAGCTTGGA<u>ATGATGCCCTTTTGCCAC</u> (SEQ ID NO: 152), and for LW1604 is:

GATCCTCGAGCA<u>TCATTCAAAGTAGGTGG</u>. (SEQ ID NO: 153). The underlined portion of the primer matches the 5' and 3' areas, respectively, of a portion of the coding region of nGPCR-56.

The PCR reaction for the coding region was performed in a 50 μl reaction containing 32 μl H₂O, 5 μl 10X TT buffer (140 mM Ammonium Sulfate, 0.1 % gelatin, 0.6 M Tris-tricine pH 8.4), 5 μl 15 mM MgS0₄, 2 μl 10 mM dNTP, 5 μl human genomic DNA (0.3μg/μl)(Clontech), 0.3 μl LW1603 (1 μg/μl)(SEQ ID NO: 152), 0.3 μl LW1696 (1 μg/μl)(SEQ ID NO: 154), 0.4 μl High Fidelity Taq polymerase (Boehringer Mannheim). The PCR reaction was started with 1 cycle of 94°C for 2 min followed by 25 cycles at 94°C for 40 sec, 55°C for 60 sec., 68°C for 2 min. The PCR reaction was loaded onto a 2 % agarose gel. The DNA band was excised from the gel, placed in GenElute Agarose spin column (Supelco) and spun for 10 min at maximum speed. The eluted DNA was EtOH precipitated and resuspended in 12μl H₂O for ligation. The PCR primer sequence for LW1603 is:

GATCAAGCTTGGA<u>ATGATGCCCTTTTGCCAC</u> (SEQ ID NO: 152), and LW1696:

GATCCTCGAGCTATGAACTCAATTCCAAAAATAATTTACACC (SEQ ID NO: 154). The underlined portion of the primer matches the 5' and 3' areas, respectively, of a portion of the coding region.

The ligation reaction used solutions from the TOPO TA Cloning Kit (Invitrogen) which consisted of 4µI PCR product DNA, 1 µI of salt solution and 1 µI pCRII-TOPO vector that was incubated for 5 minutes at room temperature then the reaction was placed on ice. Two microliters of the ligation reaction was transformed in One-Shot TOP10 cells (Invitrogen), and placed on ice for 30 minutes. The cells were heat-shocked for 30 seconds at 42°C, placed on ice for two minutes, 250 µI of SOC was added, then incubated at 37°C with shaking for one hour and then plated onto ampicillin plates. A single colony containing an insert was used to inoculate a 5 ml culture of LB medium. Plasmid DNA was purified using a Concert Rapid Plasmid Miniprep System (GibcoBRL) and then sequenced.

The mutation in nGPCR-56 was repaired using the QuikChange Site-Directed Mutagenesis Kit (Stratagene). The PCR reaction contained 40 µl H2O, 5 µl 10x

5

10

15

20

25

Reaction buffer, 1 µl mini-prep DNA, 1 µl LW1700 (125 ng/µl) (SEQ ID NO: 155), 1µl LW1701 (125 ng/µl) (SEQ ID NO: 156), 1µl 10 mM dNTP, 1 µl Pfu DNA polymerase. The cycle conditions were 95°C for 30 sec then 14 cycles at 95°C for 30 sec, 55°C for 1 min, 68°C for 12 min. The tube was placed on ice for 2 min, then 1 µl of *DpnI* was added and the tube incubated at 37°C for one hour. One microliter of the *DpnI*-treated DNA was transformed into Epicurian coli XL1-Blue supercompetent cells and the entire insert was re-sequenced. The primer sequences are: GCTACTTGAACTCTACATTTAATCCAATGGTTTATGCATTTTCTATCC (LW1700)(SEQ ID NO: 155), and:

GGATAGAAAATGCATAAACCATTGGATTAAATGTAGAGTTCAAGTAGC (LW1701)(SEQ ID NO: 156).

The DNA subcloned into pCRII was sequenced using the ABI PRISMTM 310 Genetic Analyzer (PE Applied Biosystems) which uses advanced capillary electrophoresis technology and the ABI PRISMTM BigDyeTM Terminator Cycle Sequencing Ready Reaction Kit. Each cycle-sequencing reaction contained 6 μl of H₂0, 8 μl of BigDye Terminator mix, 5 μl mini-prep DNA (0.1 μg/μl), and 1 μl primer (25 ng/μl) and was performed in a Perkin-Elmer 9600 thermocycler with 25 cycles of 96°C for 10 sec, 50°C for 10 sec, and 60°C for 4 min. The product was purified using a CentriflexTM gel filtration cartridge, dried under vacuum, then dissolved in 16 μl of Template Suppression Reagent (PE Applied Biosystems). The samples were heated at 95°C for 5 min then placed in the 310 Genetic Analyzer, yielding the sequence of SEQ ID NO: 89.

nGPCR-58: PCR AND SUBCLONING

Isolation of a clone for nGPCR-58 from genomic DNA was performed by PCR in a 50 µl reaction containing Herculase DNA Polymerse blend (Stratagene), with buffer recommendations as supplied by the manufacturer, 200 ng each primers PSK14 (SEQ ID NO: 157) and PSK15 (SEQ ID NO: 158), 150 ng of human genomic DNA (Clontech) and 6% DMSO. The PCR reaction was performed on a Robocycler thermocycler (Stratagene) starting with 1 cycle of 94°C for 2 min followed by 35 cycles of 94°C for 30 sec, 65°C for 30 sec, 72°C for 2 min. The PCR reaction was purified by the QiaQuick PCR Purification Kit (Qiagen) and eluted in TE. The PCR primer sequences were:

10

15

20

25

PSK14: 5'GATCGAATTC<u>ATGGACACTACCATGGAAGCTGACC</u> (SEQ ID NO: 157), and:

PSK15: 5'GATCCTCGAG<u>TCACGTGGGGCCTGCGCCCGG</u> (SEQ ID NO: 158).

The underlined portion of the primers match the 5' and 3' areas, respectively, of a portion of the 5' untranslated region and coding region. Translation initiation and termination codons are shown above in bold.

The blunt ended PCR product was prepared for cloning by the addition of a single base "A" residue by AmpliTaq Gold (Perkin Elmer) in a reaction with 1X PCR Buffer II, 1 mM MgCl₂, 200uM each dATP, dGTP, dCTP, and dTTP. The reaction was incubated at 94°C for 10 minutes followed by 72°C for 10 minutes. The products were cloned into the pCRII-TOPO vector (Invitrogen) using the TOPO TA cloning kit as follows: 3µ1 PCR product DNA, 1 µ1 pCRII-TOPO vector, and 1 µ1 TOPOII salt solution (1.2M NaCl, 0.06M MgCl₂) was incubated for 5 minutes at room temperature. To the ligation reaction one microliter of 6X TOPO Cloning Stop Solution was added then the reaction was placed on ice. Two microliters of the ligation reaction was transformed in One-Shot TOP10 cells (Invitrogen), and placed on ice for 30 minutes. The cells were heat-shocked for 30 seconds at 42°C, placed on ice for two minutes, 250 µl of SOC was added, then incubated at 37°C with shaking for one hour and then plated onto ampicillin plates supplemented with X-gal and IPTG. Single colonies were screened by PCR for the presence of the insert, and a plasmid DNA from colony 58-6 was purified using a Qiagen Endo-Free plasmid purification kit and deposited as nGPCR-58.

nGPCR-58 was sequenced directly using an ABI377 fluorescence-based sequencer (Perkin Elmer/Applied Biosystems Division, PE/ABD, Foster City, CA) and the ABI BigDyeTM Terminator Cycle Sequencing Ready Reaction kit with Taq FSTM polymerase. Each ABI cycle sequencing reaction contained about 0.5 μg of plasmid DNA. Cycle-sequencing was performed using an initial denaturation at 98°C for 1 min, followed by 50 cycles: 96°C for 30 sec, annealing at 50°C for 30 sec, and extension at 60°C for 4 min. Temperature cycles and times were controlled by a Perkin-Elmer 9600 thermocycler. Extension products were purified using AGTC (R) gel filtration block (Edge BiosSystems, Gaithersburg, MD). Each reaction product was loaded by pipette onto the column, which was then centrifuged in a swinging

5

10

15

20

25

bucket centrifuge (Sorvall model RT6000B tabletop centrifuge) at 1500 x g for 4 min at room temperature. Column-purified samples were dried under vacuum for about 40 min and then dissolved in 3 µl of a DNA loading solution (83% deionized formamide, 8.3 mM EDTA, and 1.6 mg/ml Blue Dextran). The samples were then heated to 90°C for 3.5 min and loaded into the gel sample wells for sequence analysis by the ABI377 sequencer. Sequence analysis was performed by importing ABI377 files into the Sequencer program (Gene Codes, Ann Arbor, MI), yielding the sequence of SEQ ID NO: 93.

EXAMPLE 3: HYBRIDIZATION ANALYSIS TO DEMONSTRATE nGPCR-X EXPRESSION IN BRAIN

The expression of nGPCR-x in mammals, such as the rat, may be investigated by *in situ* hybridization histochemistry. To investigate expression in the brain, for example, coronal and sagittal rat brain cryosections (20 µm thick) are prepared using a Reichert-Jung cryostat. Individual sections are thaw-mounted onto silanized, nuclease-free slides (CEL Associates, Inc., Houston, TX), and stored at -80°C. Sections are processed starting with post-fixation in cold 4% paraformaldehyde, rinsed in cold phosphate-buffered saline (PBS), acetylated using acetic anhydride in triethanolamine buffer, and dehydrated through a series of alcohol washes in 70%, 95%, and 100% alcohol at room temperature. Subsequently, sections are delipidated in chloroform, followed by rehydration through successive exposure to 100% and 95% alcohol at room temperature. Microscope slides containing processed cryosections are allowed to air dry prior to hybridization. Other tissues may be assayed in a similar fashion.

A nGPCR-x-specific probe is generated using PCR. Following PCR amplification, the fragment is digested with restriction enzymes and cloned into pBluescript II cleaved with the same enzymes. For production of a probe specific for the sense strand of nGPCR-x, the nGPCR-x clone in pBluescript II is linearized with a suitable restriction enzyme, which provides a substrate for labeled run-off transcripts (*i.e.*, cRNA riboprobes) using the vector-borne T7 promoter and commercially available T7 RNA polymerase. A probe specific for the antisense strand of nGPCR-x is also readily prepared using the nGPCR-x clone in pBluescript II by cleaving the recombinant plasmid with a suitable restriction enzyme to generate a linearized substrate for the production of labeled run-off cRNA transcripts using the T3

5

10

15

20

25

promoter and cognate polymerase. The riboprobes are labeled with [35 S]-UTP to yield a specific activity of about 0.40 x 10 6 cpm/pmol for antisense riboprobes and about 0.65 x 10 6 cpm/pmol for sense-strand riboprobes. Each riboprobe is subsequently denatured and added (2 pmol/ml) to hybridization buffer which contained 50% formamide, 10% dextran, 0.3 M NaCl, 10 mM Tris (pH 8.0), 1 mM EDTA, 1X Denhardt's Solution, and 10 mM dithiothreitol. Microscope slides containing sequential brain cryosections are independently exposed to 45 μ l of hybridization solution per slide and silanized cover slips are placed over the sections being exposed to hybridization solution. Sections are incubated overnight (15-18 hours) at 52°C to allow hybridization to occur. Equivalent series of cryosections are exposed to sense or antisense nGPCR-40-specific cRNA riboprobes.

Following the hybridization period, coverslips are washed off the slides in 1X SSC, followed by RNase A treatment involving the exposure of slides to 20 µg/ml RNase A in a buffer containing 10 mM Tris-HCl (pH 7.4), 0.5 M EDTA, and 0.5 M NaCl for 45 minutes at 37°C. The cryosections are then subjected to three highstringency washes in 0.1 X SSC at 52°C for 20 minutes each. Following the series of washes, cryosections are dehydrated by consecutive exposure to 70%, 95%, and 100% ammonium acetate in alcohol, followed by air drying and exposure to Kodak BioMaxTM MR-1 film. After 13 days of exposure, the film is developed. Based on these results, slides containing tissue that hybridized, as shown by film autoradiograms, are coated with Kodak NTB-2 nuclear track emulsion and the slides are stored in the dark for 32 days. The slides are then developed and counterstained with hematoxylin. Emulsion-coated sections are analyzed microscopically to determine the specificity of labeling. The signal is determined to be specific if autoradiographic grains (generated by antisense probe hybridization) are clearly associated with cresyl violate-stained cell bodies. Autoradiographic grains found between cell bodies indicates non-specific binding of the probe.

Expression of nGPCR-x in the brain provides an indication that modulators of nGPCR-x activity have utility for treating neurological disorders, including but not limited to, schizophrenia, affective disorders, ADHD/ADD (i.e., Attention Deficit-Hyperactivity Disorder/Attention Deficit Disorder), and neural disorders such as Alzheimer's disease, Parkinson's disease, migraine, and senile dementia. Some other diseases for which modulators of nGPCR-x may have utility include depression,

5

10

15

20

25

anxiety, bipolar disease, epilepsy, neuritis, neurasthenia, neuropathy, neuroses, and the like. Use of nGPCR-x modulators, including nGPCR-x ligands and anti-nGPCR-x antibodies, to treat individuals having such disease states is intended as an aspect of the invention.

EXAMPLE 4: TISSUE EXPRESSION PROFILING

Tissue specific expression of the cDNAs encoding nGPCR-1, nGPCR-3, nGPCR-9, nGPCR-11, nGPCR-16, nGPCR-40, nGPCR-54, nGPCR-56, and nGPCR-58 was detected using a PCR-based system. Tissue specific expression of cDNAs encoding nGPCR-x may be accomplished using similar methods.

Primers were synthesized by Genosys Corp., The Woodlands, TX. PCR reactions were assembled using the components of the Expand Hi-Fi PCR SystemTM (Roche Molecular Biochemicals, Indianapolis, IN).

nGPCR-1

5

10

15

20

25

30

The RapidScanTM Gene Expression Panel was used to generate a comprehensive expression profile of the putative GPCR in human tissues. Human tissues in the array may include: brain, heart, kidney, spleen, liver, colon, lung, small intestine, muscle, stomach, testis, placenta, salivary gland, thyroid, adrenal gland, pancreas, ovary, uterus, prostate, skin, PBL, bone marrow, fetal brain, fetal liver. Human brain regions in the array may include: frontal lobe, temporal lobe, cerebellum, hippocampus, substantia nigra, caudate nucleus, amygdala, thalamus, hypothalamus, pons, medulla and spinal cord.

Expression of the nGPCR-1 in the various tissues was detected by using PCR primers designed based on the available sequence of the receptor that will prime the synthesis of a 212bp fragment in the presence of the appropriate cDNA. The forward primer was:

GCTCAACCCACTCATCTATGCC (SEQ ID NO: 97), and the reverse primer was:

AAACTTCTCTGCCCTTACCGTC (SEQ ID NO: 98)

The PCR reaction mixture was added to each well of the PCR plate. The plate was placed in a GeneAmp PCR9700 PCR thermocycler (Perkin Elmer Applied Biosystems). The plate was then exposed to the following cycling parameters: Presoak 94°C for 3 min; denaturation at 94°C for 30 seconds; annealing at primer T_m for

45 seconds; extension 72°C for 2 minutes; for 35 cycles. PCR products were then separated and analyzed by electrophoresis on a 1.5-% agarose gel.

The 4-log dilution range of cDNA deposited on the plate ensured that the amplification reaction is within the linear range and, hence, facilitated the semi-quantitative determination of relative mRNA accumulation in the various tissues or brain regions examined.

Expression of nGPCR-1 was found to be highest in the testis, adrenal gland and heart. Significant levels of expression were also found in the brain, kidney, spleen ovary, prostate, muscle, PBL, stomach and bone marrow. Within the brain, expression levels were highest in the cerebellum, amygdala, thalamus and spinal cord, with significant levels of expression in the frontal lobe, hippocampus, substantia nigra, hypothalamus and pons.

Expression of nGPCR-1 in the brain provided an indication that modulators of nGPCR-1 activity have utility for treating neurological disorders, including but not limited to, schizophrenia, affective disorders, ADHD/ADD (i.e., Attention Deficit-Hyperactivity Disorder/Attention Deficit Disorder), and neural disorders such as Alzheimer's disease, Parkinson's disease, migraine, and senile dementia. Some other diseases for which modulators of nGPCR-1 may have utility include depression, anxiety, bipolar disease, epilepsy, neuritis, neurasthenia, neuropathy, neuroses, and the like. Use of nGPCR-1 modulators, including nGPCR-1 ligands and anti-nGPCR-1 antibodies, to treat individuals having such disease states is intended as an aspect of the invention.

nGPCR-3

5

10

15

20

25

30

Tissue specific expression of the cDNA encoding nGPCR-3 was detected using a PCR-based method. Multiple ChoiceTM first strand cDNAs (OriGene Technologies, Rockville, MD) from 6 human tissues were serially diluted over a 3-log range and arrayed into a multi-well PCR plate. This array was used to generate a comprehensive expression profile of the putative GPCR in human tissues. Human tissues arrayed included: brain, heart, kidney, peripheral blood leukocytes, lung and testis. PCR primers were designed based on the available sequence of the putative GPCR. The sequence of the forward primer used was:

5'TGCTGCTTTGTTGCGCCTAC3' (SEQ ID NO: 189), corresponding to base pairs 77 through 96 of the predicted coding sequence of nGPCR-3. The sequence of the reverse primer used was:

5'TTGGACGCCAGGAAGGTG3' (SEQ ID NO: 190), corresponding to base pairs 258 through 285 of the predicted coding sequence of nGPCR-3. This primer set primes the synthesis of a 298 base pair fragment in the presence of the appropriate cDNA. For detection of expression within brain regions, the same primer set was used with the Human Brain Rapid ScanTM Panel (OriGene Technologies, Rockville, MD). This panel represents serial dilutions over a 3 log range of first strand cDNA from the following brain regions arrayed in a 96 well format: frontal lobe, temporal lobe, cerebellum, hippocampus, substantia nigra, caudate nucleus, amygdala, thalamus, hypothalamus, pons, medulla and spinal cord. Primers were synthesized by Genosys Corp., The Woodlands, TX. PCR reactions were assembled using the components of the Expand Hi-Fi PCR SystemTM (Roche Molecular Biochemicals. Indianapolis, IN). Twenty-five microliters of the PCR reaction mixture was added to each well of the RapidScan PCR plate. The plate was placed in a GeneAmp 9700 PCR thermocycler (Perkin Elmer Applied Biosystems). The following cycling program was executed: Pre-soak at (94°C for 3min.) followed by 35 cycles of [(94°C for 45 sec.), (53°C for 2 min.), and (72°C for 45 sec.)]. PCR reaction products were then separated and analyzed by electrophoresis on a 2.0% agarose gel stained with ethidium bromide.

The results indicated that nGPCR-3 was expressed in the brain, heart, kidney, peripheral blood lymphocytes, lung, and testis. In the brain, nGPCR-3 was expressed in frontal lobe, temporal lobe, cerebellum, hippocampus, substantia nigra, caudate nucleus, amygdala, thalamus, hypothalamus, pons, medulla, as well as in the spinal cord.

nGPCR-9

10

15

20

25

30

The RapidScanTM Gene Expression Panel was used to generate a comprehensive expression profile of the putative GPCR in human tissues. Human tissues arrayed include: brain, heart, kidney, spleen, liver, colon, lung, small intestine, muscle, stomach, testis, placenta, salivary gland, thyroid, adrenal gland, pancreas, ovary, uterus, prostate, skin, PBL, bone marrow, fetal brain, fetal liver.

The forward primer used was to detect expression of nGPCR-9 was: 5' AACCCCATCATCTACACGC 3'(SEQ ID NO: 105), and, the reverse primer was:

5' TGCCTGTGGAGCCGCTGG 3'(SEQ ID NO: 106). This primer set will prime the synthesis of a 238 base pair fragment in the presence of the appropriate cDNA.

For detection of expression within brain regions, the same primer set was used with the Human Brain Rapid ScanTM Panel (OriGene Technologies, Rockville, MD). This panel represents serial dilutions over a 2-log range of first strand cDNA from the following brain regions arrayed in a 96 well format: frontal lobe, temporal lobe, cerebellum, hippocampus, substantia nigra, caudate nucleus, amygdala, thalamus, hypothalamus, pons, medulla and spinal cord.

Twenty-five microliters of the PCR reaction mixture was added to each well of the PCR plate. The plate was placed in a GeneAmp 9700 PCR thermocycler (Perkin Elmer Applied Biosystems). The following cycling program was executed: Pre-soak at (94°Cfor 3 min.) followed by 35 cycles of [(94°Cfor 45 sec.) (52°C for 2 min.) (72°Cfor 45 sec.)]. PCR reaction products were then separated and analyzed by electrophoresis on a 2.0% agarose gel and stained with ethidium bromide.

nGPCR-9 was expressed in the brain, peripheral blood leukocytes, heart, kidney, adrenal gland, spleen, pancreas, liver, lung, skin, bone marrow, testis, placenta, salivary gland, uterus, small intestine, muscle, stomach, and fetal liver. Within the brain, nGPCR-9 was expressed in all areas examined including the frontal lobe, temporal lobe, cerebellum, hippocampus, substantia nigra, caudate nucleus, amygdala, thalamus, hypothalamus, pons, medulla and spinal cord.

Expression of nGPCR-9 in the brain provided an indication that modulators of nGPCR-9 activity have utility for treating disorders, including but not limited to, schizophrenia, affective disorders, movement disorders, metabolic disorders, inflammatory disorders, cancers, ADHD/ADD (i.e., Attention Deficit-Hyperactivity Disorder/Attention Deficit Disorder), and neural disorders such as Alzheimer's disease, Parkinson's disease, migraine, and senile dementia. Use of nGPCR-9 modulators, including nGPCR-9 ligands and anti-nGPCR-9 antibodies, to treat individuals having such disease states is intended as an aspect of the invention.

nGPCR-11

The RapidScanTM Gene Expression Panel was used to generate a comprehensive expression profile of the putative GPCR in human tissues. Human tissues in the array included, *inter alia*: brain, heart, kidney, spleen, liver, colon, lung, small intestine, muscle, stomach, testis, placenta, salivary gland, thyroid, adrenal

5

10

15

20

25

gland, pancreas, ovary, uterus, prostate, skin, PBL, bone marrow, fetal brain, fetal liver. Human brain regions in the array included, *inter alia*: frontal lobe, temporal lobe, cerebellum, hippocampus, substantia nigra, caudate nucleus, amygdala, thalamus, hypothalamus, pons, medulla and spinal cord.

Expression of nGPCR-11 in the various tissues was detected by using PCR primers designed based on the available sequence of the receptor that will prime the synthesis of a 206bp fragment in the presence of the appropriate cDNA. The forward primer used to detect expression of nGPCR-11 was:

5'-GAAGCCCAGCACTGTTTACC-3' (SEQ ID NO: 109), and the reverse primer was:

5'-TGAAATACCTGTCCGCAGCC-3 (SEQ ID NO: 110).

Twenty-five microliters of the PCR reaction mixture was added to each well of the RapidScan PCR plate. The plate was placed in a GeneAmp 9700 PCR thermocycler (PE Applied Biosystems). The following cycling program was executed: Pre-soak 94°C for 3 min; denaturation at 94°C for 30 seconds; annealing at primer T_m for 45 seconds; extension at 72°C for 2 minutes; for 35 cycles. PCR reaction products were then separated and analyzed by electrophoresis on a 2.0% agarose gel stained with ethidium bromide.

The 4-log dilution range of cDNA deposited on the plate ensured that the amplification reaction was within the linear range and, facilitated semi-quantitative determination of relative mRNA accumulation in the various tissues or brain regions examined.

nGPCR-11 was expressed in the thyroid gland, brain, heart, kidney, adrenal gland, spleen, liver, ovary, muscle, testis, salivary gland, colon, prostate, small intestine, skin stomach, bone marrow, fetal brain and placenta. Within the brain, nGPCR-11 was expressed in the temporal lobe, amygdala, substantia nigra, pons, spinal cord, frontal lobe, and cerebellum.

Expression of the nGPCR-11 in the brain provided an indication that modulators of nGPCR-11 activity have utility for treating disorders, including but not limited to, schizophrenia, affective disorders, metabolic disorders, inflammatory disorders, cancers, ADHD/ADD (i.e., Attention Deficit-Hyperactivity Disorder/Attention Deficit Disorder), and neural disorders such as Alzheimer's disease, Parkinson's disease, migraine, and senile dementia. Some other diseases for which modulators of nGPCR-11 may have utility include depression, anxiety, bipolar

5

10

15

20

25

disease, epilepsy, neuritis, neurasthenia, neuropathy, neuroses, and the like. Use of nGPCR-11 modulators, including nGPCR-11 ligands and anti-nGPCR-11 antibodies, to treat individuals having such disease states is intended as an aspect of the invention.

Expression of nGPCR-11 in the thyroid gland, indicates that agonists or antagonists could be of use in the treatment of thyroid dysfunction such as thyreotoxicosis and myxoedema. They could also be of use in the stimulation of thyroid hormone release leading to overall increase in metabolic rate and weight reduction. The expression of nGPCR-11 in liver and muscle indicate a use for agonists or antagonists in regulation of glucose metabolism applicable in diabetes type II.

nGCPR-16

5

10

15

20

25

30

The RapidScanTM Gene Expression Panel was used to generate a comprehensive expression profile of the putative GPCR in human tissues. Human tissues in the array included, *inter alia*: brain, heart, kidney, spleen, liver, colon, lung, small intestine, muscle, stomach, testis, placenta, salivary gland, thyroid, adrenal gland, pancreas, ovary, uterus, prostate, skin, PBL, bone marrow, fetal brain, fetal liver. Human brain regions in the array included, *inter alia*: frontal lobe, temporal lobe, cerebellum, hippocampus, substantia nigra, caudate nucleus, amygdala, thalamus, hypothalamus, pons, medulla and spinal cord.

Expression of nGPCR-16 in the various tissues was detected by using PCR primers designed based on the available sequence of the receptor that will prime the synthesis of a 205bp fragment in the presence of the appropriate cDNA. The forward primer used to detect expression of nGPCR-16 was:

5' CAGCCCAAACATCCAAGTC 3'. (SEQ ID NO: 113). The reverse primer used to detect expression of nGPCR-16 was:

5' ACCCCACTTAATCAGCCTC 3'(SEQ ID NO: 114).

For detection of expression within brain regions, the same primer set was used with the Human Brain Rapid ScanTM Panel (OriGene Technologies, Rockville, MD). This panel represents serial dilutions over a 2 log range of first strand cDNA from the following brain regions arrayed in a 96 well format: frontal lobe, temporal lobe, cerebellum, hippocampus, substantia nigra, caudate nucleus, amygdala, thalamus, hypothalamus, pons, medulla and spinal cord.

Twenty-five microliters of the PCR reaction mixture was added to each well of the RapidScan PCR plate. The plate was placed in a GeneAmp 9700 PCR thermocycler (Perkin Elmer Applied Biosystems). The following cycling program was executed: Pre-soak at (94° for 3min.) followed by 35 cycles of [(94°C for 45 sec.) (53°C for 2 min.) (72°C for 45 sec.)]. PCR reaction products were then separated and analyzed by electrophoresis on a 2.0% agarose gel, and stained with ethidium bromide.

The 4-log dilution range of cDNA deposited on the plate ensured that the amplification reaction was within the linear range and, facilitated semi-quantitative determination of relative mRNA accumulation in the various tissues or brain regions examined.

nGPCR-16 was expressed in the ovary, lung, prostate, bone marrow, salivary gland, heart, adrenal gland, spleen, liver, small intestine, skin, muscle, peripheral blood leukocytes, testis, placenta, fetal liver, brain, thyroid gland, kidney, pancreas, colon, uterus, and stomach.. Within the brain, nGPCR-16 was expressed in all areas examined including the frontal lobe, temporal lobe, cerebellum, hippocampus, substantia nigra, caudate nucleus, amygdala, thalamus, hypothalamus, pons, medulla and spinal cord.

Expression of nGPCR-16 in the brain provides an indication that modulators of nGPCR-16 activity have utility for treating neurological disorders, including but not limited to, schizophrenia, affective disorders, ADHD/ADD (*i.e.*, Attention Deficit-Hyperactivity Disorder/Attention Deficit Disorder), and neural disorders such as Alzheimer's disease, Parkinson's disease, migraine, and senile dementia. Some other diseases for which modulators of nGPCR-16 may have utility include depression, anxiety, bipolar disease, epilepsy, neuritis, neurasthenia, neuropathy, neuroses, and the like. Use of nGPCR-16 modulators, including nGPCR-16 ligands and anti-nGPCR-16 antibodies, to treat individuals having such disease states is intended as an aspect of the invention.

nGPCR-40

The RapidScanTM Gene Expression Panel (OriGene Technologies, Rockville, MD) was used to generate a comprehensive expression profile of the putative GPCR in human tissues. Human tissues arrayed include: brain, heart, kidney, spleen, liver, colon, lung, small intestine, muscle, stomach, testis, placenta, salivary gland, thyroid,

5

10

15

20

25

adrenal gland, pancreas, ovary, uterus, prostate, skin, PBL, bone marrow, fetal brain, fetal liver. The forward primer used was:

5'ACAGCCCAAAGCCAAACAC3', (SEQ ID NO: 117), and the reverse primer was:

5'CCGCAGGAGCAATGAAAATCAG3', (SEQ ID NO: 118). This primer set primed the synthesis of a 220 base pair fragment in the presence of the appropriate cDNA. For detection of expression within brain regions, the same primer set was used with the Human Brain RapidScanTM Panel (OriGene Technologies, Rockville, MD). This panel represents serial dilutions over a 2 log range of first strand cDNA from the following brain regions arrayed in a 96 well format: frontal lobe, temporal lobe, cerebellum, hippocampus, substantia nigra, caudate nucleus, amygdala, thalamus, hypothalamus, pons, medulla and spinal cord.

Twenty-five microliters of the PCR reaction mixture was added to each well of the RapidScan PCR plate. The plate was placed in a GeneAmp 9700 PCR thermocycler (Perkin Elmer Applied Biosystems). The following cycling program was executed: Pre-soak at (94°C for 3min.) followed by 35 cycles of [(94° for 45 sec.) (54°C for 2 min.) (72° for 45 sec.)]. PCR reaction products were then separated and analyzed by electrophoresis on a 2.0% agarose gel stained with ethidium bromide.

The dilution range of cDNA deposited on the plates ensured that the amplification reaction was within the linear range and, hence, facilitated semi-quantitative determination of relative mRNA accumulation in the various tissues or brain regions examined.

nGPCR-40 was expressed in the brain, peripheral blood lymphocytes, pancreas, ovary, uterus, testis, salivary gland, kidney, adrenal gland, liver, bone marrow, prostate, fetal liver, colon, muscle, and fetal brain, may be found in many other tissues, including, but not limited to, lung, small intestine, fetal brain cord, and bone. Within the brain, nGPCR-40 was expressed in the frontal lobe, hypothalamus, pons, cerebellum, caudate nucleus, and medulla.

Expression of nGPCR-40 in the brain provides an indication that modulators of nGPCR-40 activity have utility for treating neurological disorders, including but not limited to, movement disorders, affective disorders, metabolic disorders, inflammatory disorders and cancers. Use of nGPCR-40 modulators, including nGPCR-40 ligands and anti-nGPCR-40 antibodies, to treat individuals having such disease states is intended as an aspect of the invention.

5

10

15

20

25

nGPCR-54

5

10

15

20

25

30

SNST DOI: NWO 0136473A2 114

Multiple ChoiceTM first strand cDNAs (OriGene Technologies, Rockville, MD) from 12 human tissues were serially diluted over a 3-log range and arrayed into a multi-well PCR plate. Human tissues arrayed include: brain, heart, kidney, peripheral blood leukocytes, liver, lung, muscle, ovary, prostate, small intestine, spleen and testis. PCR primers were designed based on the sequence of nGPCR-54 provided herein. The forward primer used was:

5'CTGTCTCTGTCCTCTTCC3',(SEQ ID NO: 123). The reverse primer used was:

5'GCACCGATCTTCATTGAATTTC3',(SEQ ID NO: 124). This primer set primes the synthesis of a 145 base pair fragment in the presence of the appropriate cDNA. For detection of expression within brain regions, the same primer set was used with the Human Brain Rapid ScanTM Panel (OriGene Technologies, Rockville, MD). This panel represents serial dilutions over a 3 log range of first strand cDNA from the following brain regions arrayed in a 96 well format: frontal lobe, temporal lobe, cerebellum, hippocampus, substantia nigra, caudate nucleus, amygdala, thalamus, hypothalamus, pons, medulla and spinal cord.

Twenty-five microliters of the PCR reaction mixture was added to each well of the RapidScan PCR plate. The plate was placed in a GeneAmp 9700 PCR thermocycler (Perkin Elmer Applied Biosystems). The following cycling program was executed: Pre-soak at (94°C for 3min.) followed by 35 cycles of [(94°C for 45 sec.) (52.5°C for 2 min.) (72°C for 45 sec.)]. PCR reaction products were then separated and analyzed by electrophoresis on a 2.0% agarose gel stained with ethidium bromide.

nGPCR-54 was expressed in the brain, kidney, lung, muscle, testis, heart, liver, ovary, prostate, small intestine, spleen, and peripheral blood leukocytes. Within the brain, nGPCR-54 was expressed in the cerebellum, hippocampus, substantia nigra, thalamus, hypothalamus, pons, frontal lobe, temporal lobe, caudate nucleus, medulla, spinal cord, and amygdala.

Expression of the nGPCR-54 in the brain provides an indication that modulators of nGPCR-54 activity have utility for treating neurological disorders, including but not limited to, movement disorders, affective disorders, metabolic disorders, inflammatory disorders and cancers. Use of nGPCR-54 modulators,

including nGPCR-54 ligands and anti-nGPCR-54 antibodies, to treat individuals having such disease states is intended as an aspect of the invention.

nGPCR-56

5

10

15

20

25

30

The RapidScanTM Gene Expression Panel was used to generate a comprehensive expression profile of the putative GPCR in human tissues. Human tissues arrayed include: brain, heart, kidney, spleen, liver, colon, lung, small intestine, muscle, stomach, testis, placenta, salivary gland, thyroid, adrenal gland, pancreas, ovary, uterus, prostate, skin, PBL, bone marrow, fetal brain, fetal liver. The forward primer used was:

5' ACTTCAAACAACTTCATACCCC 3' (SEQ ID NO: 125), and the reverse primer used was:

5'ACACACAGCATAGTAGCG 3' (SEQ ID NO: 126). This primer set will prime the synthesis of a 231 base pair fragment in the presence of the appropriate cDNA. For detection of expression within brain regions, the same primer set was used with the Human Brain Rapid ScanTM Panel (OriGene Technologies, Rockville, MD). This panel represents serial dilutions over a 2 log range of first strand cDNA from the following brain regions arrayed in a 96 well format: frontal lobe, temporal lobe, cerebellum, hippocampus, substantia nigra, caudate nucleus, amygdala, thalamus, hypothalamus, pons, medulla and spinal cord.

Twenty-five microliters of the PCR reaction mixture was added to each well of the RapidScan PCR plate. The plate was placed in a GeneAmp 9700 PCR thermocycler (Perkin Elmer Applied Biosystems). The following cycling program was executed: Pre-soak at (94°C for 3min.) followed by 35 cycles of [(94°C for 45 sec.) (53°C for 2 min.) (72°C for 45 sec.)]. PCR reaction products were then separated and analyzed by electrophoresis on a 2.0% agarose gel stained with ethidium bromide.

nGPCR-56 was expressed in peripheral blood lymphocytes, testis, salivary gland, kidney, spleen, skin, stomach, placenta, ovary, bone marrow, fetal liver, small intestine, and fetal brain.

Expression of nGPCR-56 in the brain provides an indication that modulators of nGPCR-56 activity have utility for treating neurological disorders, including but not limited to, movement disorders, affective disorders, metabolic disorders, inflammatory disorders and cancers. Use of nGPCR-56 modulators, including

nGPCR-56 ligands and anti-nGPCR-56 antibodies, to treat individuals having such disease states is intended as an aspect of the invention.

nGPCR-58

5

10

15

20

25

30

The RapidScanTM Gene Expression Panel was used to generate a comprehensive expression profile of the putative GPCR in human tissues. Human tissues in the array included: brain, heart, kidney, spleen, liver, lung, small intestine, muscle, testis, ovary, prostate, and PBL. Human brain regions in the array included: frontal lobe, temporal lobe, cerebellum, hippocampus, substantia nigra, caudate nucleus, amygdala, thalamus, hypothalamus, pons, medulla and spinal cord.

Expression of the nGPCR-58 in the various tissues was detected by using PCR primers designed based on the available sequence of the receptor that will prime the synthesis of a 282bp fragment in the presence of the appropriate cDNA. The forward primer was:

CCCATAGGAAGTAGTAGAAG (SEQ ID NO: 128).

The PCR reaction mixture was added to each well of the PCR plate. The plate was placed in a GeneAmp PCR9700 PCR thermocycler (Perkin Elmer Applied Biosystems). The plate was then exposed to the following cycling parameters: Presoak 94° for 3 min; denaturation at 94° for 30 seconds; annealing at primer T_m for 45 seconds; extension at 72° for 2 minutes; for 35 cycles. PCR productions were then separated and analyzed by electrophoresis on a 1.5-% agarose gel.

The 4-log dilution range of cDNA deposited on the plate ensured that the amplification reaction was within the linear range and, hence, facilitated semi-quantitative determination of relative mRNA accumulation in the various tissues or brain regions examined.

nGPCR-58 was expressed in all tissues included on the array, including brain, muscle, prostate, kidney, peripheral blood lymphocytes, liver, lung, small intestine, spleen, testis, heart, and ovary. Within the brain, nGPCR-58 was expressed in many regions including, but not limited to cerebellum, substantia nigra, thalamus, pons, spinal cord, frontal lobe, temporal lobe, hippocampus, caudate nucleus, amygdala, hypothalamus, and medulla.

Expression of the nGPCR-58 in the brain provided an indication that modulators of nGPCR-58 activity have utility for treating disorders, including but not

limited to, schizophrenia, affective disorders, ADHD/ADD (*i.e.*, Attention Deficit-Hyperactivity Disorder/Attention Deficit Disorder), neural disorders such as Alzheimer's disease, Parkinson's disease, migraine, senile dementia, depression, anxiety, bipolar disease, epilepsy, neuritis, neurasthenia, neuropathy, neuroses, metabolic disorders, inflammatory disorders, cancers and the like. Use of nGPCR-58 modulators, including nGPCR-58 ligands and anti-nGPCR-58 antibodies, to treat individuals having such disease states is intended as an aspect of the invention.

EXAMPLE 5: NORTHERN BLOT ANALYSIS

Northern blots are performed to examine the expression of nGPCR-x mRNA. The sense orientation oligonucleotide and the antisense-orientation oligonucleotide, described above, are used as primers to amplify a portion of the GPCR-x cDNA sequence of an odd numbered nucleotide sequence ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185.

Multiple human tissue northern blots from Clontech (Human II # 7767-1) are hybridized with the probe. Pre-hybridization is carried out at 42 C for 4 hours in 5xSSC, 1X Denhardt's reagent, 0.1% SDS, 50% formamide, 250 mg/ml salmon sperm DNA. Hybridization is performed overnight at 42°C in the same mixture with the addition of about $1.5x10^6$ cpm/ml of labeled probe.

The probe is labeled with α-³²P-dCTP by RediprimeTM DNA labeling system (Amersham Pharmacia), purified on Nick ColumnTM (Amersham Pharmacia) and added to the hybridization solution. The filters are washed several times at 42 C in 0.2x SSC, 0.1% SDS. Filters are exposed to Kodak XAR film (Eastman Kodak Company, Rochester, N.Y., USA) with intensifying screen at –80°C.

EXAMPLE 6: RECOMBINANT EXPRESSION OF nGPCR-X IN EUKARYOTIC HOST CELLS

A. Expression of nGPCR-x in Mammalian Cells

To produce nGPCR-x protein, a nGPCR-x-encoding polynucleotide is expressed in a suitable host cell using a suitable expression vector and standard genetic engineering techniques. For example, the nGPCR-x-encoding sequence described in Example 1 is subcloned into the commercial expression vector pzeoSV2 (Invitrogen, San Diego, CA) and transfected into Chinese Hamster Ovary (CHO) cells

10

15

20

25

using the transfection reagent FuGENE6TM (Boehringer-Mannheim) and the transfection protocol provided in the product insert. Other eukaryotic cell lines, including human embryonic kidney (HEK 293) and COS cells, are suitable as well. Cells stably expressing nGPCR-x are selected by growth in the presence of 100 μg/ml zeocin (Stratagene, LaJolla, CA). Optionally, nGPCR-x may be purified from the cells using standard chromatographic techniques. To facilitate purification, antisera is raised against one or more synthetic peptide sequences that correspond to portions of the nGPCR-x amino acid sequence, and the antisera is used to affinity purify nGPCR-x. The nGPCR-x also may be expressed in-frame with a tag sequence (e.g., polyhistidine, hemagluttinin, FLAG) to facilitate purification. Moreover, it will be appreciated that many of the uses for nGPCR-x polypeptides, such as assays described below, do not require purification of nGPCR-x from the host cell.

B. Expression of nGPCR-x in 293 cells

5

10

15

20

25

30

BNSDGG D RWIT 1156475AT

For expression of nGPCR-x in mammalian cells 293 (transformed human, primary embryonic kidney cells), a plasmid bearing the relevant nGPCR-x coding sequence is prepared, using vector pSecTag2A (Invitrogen). Vector pSecTag2A contains the murine IgK chain leader sequence for secretion, the c-myc epitope for detection of the recombinant protein with the anti-myc antibody, a C-terminal polyhistidine for purification with nickel chelate chromatography, and a Zeocin resistant gene for selection of stable transfectants. The forward primer for amplification of this GPCR cDNA is determined by routine procedures and preferably contains a 5' extension of nucleotides to introduce the *HindIII* cloning site and nucleotides matching the GPCR sequence. The reverse primer is also determined by routine procedures and preferably contains a 5' extension of nucleotides to introduce an *XhoI* restriction site for cloning and nucleotides corresponding to the reverse complement of the nGPCR-x sequence. The PCR conditions are 55°C as the annealing temperature. The PCR product is gel purified and cloned into the *HindIII-XhoI* sites of the vector.

The DNA is purified using Qiagen chromatography columns and transfected into 293 cells using DOTAPTM transfection media (Boehringer Mannheim, Indianapolis, IN). Transiently transfected cells are tested for expression after 24 hours of transfection, using western blots probed with anti-His and anti-nGPCR-x

peptide antibodies. Permanently transfected cells are selected with Zeocin and propagated. Production of the recombinant protein is detected from both cells and media by western blots probed with anti-His, anti-Myc or anti-GPCR peptide antibodies.

C. Expression of nGPCR-x in COS cells

For expression of the nGPCR-x in COS7 cells, a polynucleotide molecule having an odd numbered nucleotide sequence ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185 can be cloned into vector p3-CI. This vector is a pUC18-derived plasmid that contains the HCMV (human cytomegalovirus) promoter-intron located upstream from the bGH (bovine growth hormone) polyadenylation sequence and a multiple cloning site. In addition, the plasmid contains the dhrf (dihydrofolate reductase) gene which provides selection in the presence of the drug methotrexane (MTX) for selection of stable transformants.

The forward primer is determined by routine procedures and preferably contains a 5' extension which introduces an XbaI restriction site for cloning, followed by nucleotides which correspond to an odd numbered nucleotide sequence ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185. The reverse primer is also determined by routine procedures and preferably contains 5'- extension of nucleotides which introduces a Sall cloning site followed by nucleotides which correspond to the reverse complement of an odd numbered nucleotide sequence ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185. The PCR consists of an initial denaturation step of 5 min at 95°C 30 cycles of 30 sec denaturation at 95°C, 30 sec annealing at 58°C and 30 sec extension at 72°C, followed by 5 min extension at 72°C. The PCR product is gel purified and ligated into the XbaI and SalI sites of vector p3-CI. This construct is transformed into E. coli cells for amplification and DNA purification. The DNA is purified with Qiagen chromatography columns and transfected into COS 7 cells using Lipofectamine™ reagent from BRL, following the manufacturer's protocols. Forty-eight and 72 hours after transfection, the media and the cells are tested for recombinant protein expression.

nGPCR-x expressed from a COS cell culture can be purified by concentrating the cell-growth media to about 10 mg of protein/ml, and purifying the protein by, for

5

10

15

20

25

example, chromatography. Purified nGPCR-x is concentrated to 0.5 mg/ml in an Amicon concentrator fitted with a YM-10 membrane and stored at -80°C.

D. Expression of nGPCR-x in Insect Cells

5

10

15

20

25

30

7936473A3 S

BNSDGG TELWIS

For expression of nGPCR-x in a baculovirus system, a polynucleotide molecule having an odd numbered nucleotide sequence ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185 can be amplified by PCR. The forward primer is determined by routine procedures and preferably contains a 5' extension which adds the *NdeI* cloning site, followed by nucleotides which correspond to an odd numbered nucleotide sequence ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185. The reverse primer is also determined by routine procedures and preferably contains a 5' extension which introduces the *KpnI* cloning site, followed by nucleotides which correspond to the reverse complement of an odd numbered nucleotide sequence ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185.

The PCR product is gel purified, digested with *NdeI* and *KpnI*, and cloned into the corresponding sites of vector pACHTL-A (Pharmingen, San Diego, CA). The pAcHTL expression vector contains the strong polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV), and a 6XHis tag upstream from the multiple cloning site. A protein kinase site for phosphorylation and a thrombin site for excision of the recombinant protein precede the multiple cloning site is also present. Of course, many other baculovirus vectors could be used in place of pAcHTL-A, such as pAc373, pVL941 and pAcIM1. Other suitable vectors for the expression of GPCR polypeptides can be used, provided that the vector construct includes appropriately located signals for transcription, translation, and trafficking, such as an in-frame AUG and a signal peptide, as required. Such vectors are described in Luckow *et al.*, Virology 170:31-39, among others.

The virus is grown and isolated using standard baculovirus expression methods, such as those described in Summers *et al.* (A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures, Texas Agricultural Experimental Station Bulletin No. 1555 (1987)).

In a preferred embodiment, pAcHLT-A containing nGPCR-x gene is introduced into baculovirus using the "BaculoGoldTM" transfection kit (Pharmingen,

San Diego, CA) using methods established by the manufacturer. Individual virus isolates are analyzed for protein production by radiolabeling infected cells with ³⁵S-methionine at 24 hours post infection. Infected cells are harvested at 48 hours post infection, and the labeled proteins are visualized by SDS-PAGE. Viruses exhibiting high expression levels can be isolated and used for scaled up expression.

For expression of a nGPCR-x polypeptide in a Sf9 cells, a polynucleotide molecule having the sequence of an odd numbered nucleotide sequence ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185 can be amplified by PCR using the primers and methods described above for baculovirus expression. The nGPCR-x cDNA is cloned into vector pAcHLT-A (Pharmingen) for expression in Sf9 insect. The insert is cloned into the *NdeI* and *KpnI* sites, after elimination of an internal *NdeI* site (using the same primers described above for expression in baculovirus). DNA is purified with Qiagen chromatography columns and expressed in Sf9 cells. Preliminary Western blot experiments from non-purified plaques are tested for the presence of the recombinant protein of the expected size which reacted with the GPCR-specific antibody. These results are confirmed after further purification and expression optimization in HiG5 cells.

EXAMPLE 7: INTERACTION TRAP/TWO-HYBRID SYSTEM

In order to assay for nGPCR-x-interacting proteins, the interaction trap/two-hybrid library screening method can be used. This assay was first described in Fields et al., Nature, 1989, 340, 245, which is incorporated herein by reference in its entirety. A protocol is published in Current Protocols in Molecular Biology 1999, John Wiley & Sons, NY, and Ausubel, F. M. et al. 1992, Short protocols in molecular biology, Fourth edition, Greene and Wiley-interscience, NY, each of which is incorporated herein by reference in its entirety. Kits are available from Clontech, Palo Alto, CA (Matchmaker Two-Hybrid System 3).

A fusion of the nucleotide sequences encoding all or partial nGPCR-x and the yeast transcription factor GAL4 DNA-binding domain (DNA-BD) is constructed in an appropriate plasmid (*i.e.*, pGBKT7) using standard subcloning techniques. Similarly, a GAL4 active domain (AD) fusion library is constructed in a second plasmid (*i.e.*, pGADT7) from cDNA of potential GPCR-binding proteins (for protocols on forming cDNA libraries, see Sambrook *et al.* 1989, Molecular cloning: a laboratory manual,

5

10

15

20

25

5

10

15

20

25

30

BASCOTT BWY HILLIAMAY

second edition, Cold Spring Harbor Press, Cold Spring Harbor, NY), which is incorporated herein by reference in its entirety. The DNA-BD/nGPCR-x fusion construct is verified by sequencing, and tested for autonomous reporter gene activation and cell toxicity, both of which would prevent a successful two-hybrid analysis. Similar controls are performed with the AD/library fusion construct to ensure expression in host cells and lack of transcriptional activity. Yeast cells are transformed (ca. 105 transformants/mg DNA) with both the nGPCR-x and library fusion plasmids according to standard procedures (Ausubel et al., 1992, Short protocols in molecular biology, fourth edition, Greene and Wiley-interscience, NY, which is incorporated herein by reference in its entirety). In vivo binding of DNA-BD/nGPCR-x with AD/library proteins results in transcription of specific yeast plasmid reporter genes (i.e., lacZ, HIS3, ADE2, LEU2). Yeast cells are plated on nutrient-deficient media to screen for expression of reporter genes. Colonies are dually assayed for β-galactosidase activity upon growth in Xgal (5-bromo-4-chloro-3indolyl-β-D-galactoside) supplemented media (filter assay for β-galactosidase activity is described in Breeden et al., Cold Spring Harb. Symp. Quant. Biol., 1985, 50, 643, which is incorporated herein by reference in its entirety). Positive AD-library plasmids are rescued from transformants and reintroduced into the original yeast strain as well as other strains containing unrelated DNA-BD fusion proteins to confirm specific nGPCR-x/library protein interactions. Insert DNA is sequenced to verify the presence of an open reading frame fused to GAL4 AD and to determine the identity of the nGPCR-x-binding protein.

EXAMPLE 8: MOBILITY SHIFT DNA-BINDING ASSAY USING GEL ELECTROPHORESIS

A gel electrophoresis mobility shift assay can rapidly detect specific protein-DNA interactions. Protocols are widely available in such manuals as Sambrook *et al.* **1989**, *Molecular cloning: a laboratory* manual, second edition, Cold Spring Harbor Press, Cold Spring Harbor, NY and Ausubel, F. M. *et al.*, 1992, Short Protocols in Molecular Biology, fourth edition, Greene and Wiley-interscience, NY, each of which is incorporated herein by reference in its entirety.

Probe DNA(<300 bp) is obtained from synthetic oligonucleotides, restriction endonuclease fragments, or PCR fragments and end-labeled with ³²P. An aliquot of

purified nGPCR-x (ca. 15 μg) or crude nGPCR-x extract (ca. 15 ng) is incubated at constant temperature (in the range 22-37 C) for at least 30 minutes in 10-15 μl of buffer (i.e. TAE or TBE, pH 8.0-8.5) containing radiolabeled probe DNA, nonspecific carrier DNA (ca. 1 μg), BSA (300 μg/ml), and 10% (v/v) glycerol. The reaction mixture is then loaded onto a polyacrylamide gel and run at 30-35 mA until good separation of free probe DNA from protein-DNA complexes occurs. The gel is then dried and bands corresponding to free DNA and protein-DNA complexes are detected by autoradiography.

EXAMPLE 9: ANTIBODIES TO nGPCR-X

Standard techniques are employed to generate polyclonal or monoclonal antibodies to the nGPCR-x receptor, and to generate useful antigen-binding fragments thereof or variants thereof, including "humanized" variants. Such protocols can be found, for example, in Sambrook *et al.* (1989) and Harlow *et al.* (Eds.), <u>Antibodies A Laboratory Manual;</u> Cold Spring Harbor Laboratory; Cold Spring Harbor, NY (1988). In one embodiment, recombinant nGPCR-x polypeptides (or cells or cell membranes containing such polypeptides) are used as antigen to generate the antibodies. In another embodiment, one or more peptides having amino acid sequences corresponding to an immunogenic portion of nGPCR-x (*e.g.*, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or more amino acids) are used as antigen. Peptides corresponding to extracellular portions of nGPCR-x, especially hydrophilic extracellular portions, are preferred. The antigen may be mixed with an adjuvant or linked to a hapten to increase antibody production.

A. Polyclonal or Monoclonal antibodies

As one exemplary protocol, recombinant nGPCR-x or a synthetic fragment thereof is used to immunize a mouse for generation of monoclonal antibodies (or larger mammal, such as a rabbit, for polyclonal antibodies). To increase antigenicity, peptides are conjugated to Keyhole Lympet Hemocyanin (Pierce), according to the manufacturer's recommendations. For an initial injection, the antigen is emulsified with Freund's Complete Adjuvant and injected subcutaneously. At intervals of two to three weeks, additional aliquots of nGPCR-x antigen are emulsified with Freund's Incomplete Adjuvant and injected subcutaneously. Prior to the final booster injection, a serum sample is taken from the immunized mice and assayed by western blot to

5

10

15

20

25

confirm the presence of antibodies that immunoreact with nGPCR-x. Serum from the immunized animals may be used as polyclonal antisera or used to isolate polyclonal antibodies that recognize nGPCR-x. Alternatively, the mice are sacrificed and their spleen removed for generation of monoclonal antibodies.

To generate monoclonal antibodies, the spleens are placed in 10 ml serum-free RPMI 1640, and single cell suspensions are formed by grinding the spleens in serum-free RPMI 1640, supplemented with 2 mM L-glutamine, 1 mM sodium pyruvate, 100 units/ml penicillin, and 100 μ g/ml streptomycin (RPMI) (Gibco, Canada). The cell suspensions are filtered and washed by centrifugation and resuspended in serum-free RPMI. Thymocytes taken from three naive Balb/c mice are prepared in a similar manner and used as a Feeder Layer. NS-1 myeloma cells, kept in log phase in RPMI with 10% fetal bovine serum (FBS) (Hyclone Laboratories, Inc., Logan, Utah) for three days prior to fusion, are centrifuged and washed as well.

To produce hybridoma fusions, spleen cells from the immunized mice are combined with NS-1 cells and centrifuged, and the supernatant is aspirated. The cell pellet is dislodged by tapping the tube, and 2 ml of 37°C PEG 1500 (50% in 75 mM HEPES, pH 8.0) (Boehringer-Mannheim) is stirred into the pellet, followed by the addition of serum-free RPMI. Thereafter, the cells are centrifuged, resuspended in RPMI containing 15% FBS, 100 μ M sodium hypoxanthine, 0.4 μ M aminopterin, 16 μ M thymidine (HAT) (Gibco), 25 units/ml IL-6 (Boehringer-Mannheim) and 1.5 x 10⁶ thymocytes/ml, and plated into 10 Corning flat-bottom 96-well tissue culture plates (Corning, Corning New York).

On days 2, 4, and 6 after the fusion, $100 \mu l$ of medium is removed from the wells of the fusion plates and replaced with fresh medium. On day 8, the fusions are screened by ELISA, testing for the presence of mouse IgG that binds to nGPCR-x. Selected fusion wells are further cloned by dilution until monoclonal cultures producing anti-nGPCR-x antibodies are obtained.

B. <u>Humanization of anti-nGPCR-x monoclonal antibodies</u>

The expression pattern of nGPCR-x as reported herein and the proven track record of GPCRs as targets for therapeutic intervention suggest therapeutic indications for nGPCR-x inhibitors (antagonists). nGPCR-x-neutralizing antibodies comprise one class of therapeutics useful as nGPCR-x antagonists. Following are

5

10

15

20

25

protocols to improve the utility of anti-nGPCR-x monoclonal antibodies as therapeutics in humans by "humanizing" the monoclonal antibodies to improve their serum half-life and render them less immunogenic in human hosts (*i.e.*, to prevent human antibody response to non-human anti-nGPCR-x antibodies).

The principles of humanization have been described in the literature and are facilitated by the modular arrangement of antibody proteins. To minimize the possibility of binding complement, a humanized antibody of the IgG4 isotype is preferred.

For example, a level of humanization is achieved by generating chimeric antibodies comprising the variable domains of non-human antibody proteins of interest with the constant domains of human antibody molecules. (See, e.g., Morrison et al., Adv. Immunol., 44:65-92 (1989)). The variable domains of nGPCR-x-neutralizing anti-nGPCR-x antibodies are cloned from the genomic DNA of a B-cell hybridoma or from cDNA generated from mRNA isolated from the hybridoma of interest. The V region gene fragments are linked to exons encoding human antibody constant domains, and the resultant construct is expressed in suitable mammalian host cells (e.g., myeloma or CHO cells).

To achieve an even greater level of humanization, only those portions of the variable region gene fragments that encode antigen-binding complementarity determining regions ("CDR") of the non-human monoclonal antibody genes are cloned into human antibody sequences. (See, e.g., Jones et al., Nature 321:522-525 (1986); Riechmann et al., Nature 332:323-327 (1988); Verhoeyen et al., Science 239:1534-36 (1988); and Tempest et al., Bio/Technology 9: 266-71 (1991)). If necessary, the β -sheet framework of the human antibody surrounding the CDR3 regions also is modified to more closely mirror the three dimensional structure of the antigen-binding domain of the original monoclonal antibody. (See Kettleborough et al., Protein Engin., 4:773-783 (1991); and Foote et al., J. Mol. Biol., 224:487-499 (1992)).

In an alternative approach, the surface of a non-human monoclonal antibody of interest is humanized by altering selected surface residues of the non-human antibody, *e.g.*, by site-directed mutagenesis, while retaining all of the interior and contacting residues of the non-human antibody. See Padlan, Molecular Immunol., 28(4/5):489-98 (1991).

5

10

15

20

25

The foregoing approaches are employed using nGPCR-x-neutralizing anti-nGPCR-x monoclonal antibodies and the hybridomas that produce them to generate humanized nGPCR-x-neutralizing antibodies useful as therapeutics to treat or palliate conditions wherein nGPCR-x expression or ligand-mediated nGPCR-x signaling is detrimental.

C. Human nGPCR-x-Neutralizing Antibodies from Phage Display

Human nGPCR-x-neutralizing antibodies are generated by phage display techniques such as those described in Aujame *et al.*, Human Antibodies 8(4):155-168 (1997); Hoogenboom, TIBTECH 15:62-70 (1997); and Rader *et al.*, Curr. Opin. Biotechnol. 8:503-508 (1997), all of which are incorporated by reference. For example, antibody variable regions in the form of Fab fragments or linked single chain Fv fragments are fused to the amino terminus of filamentous phage minor coat protein pIII. Expression of the fusion protein and incorporation thereof into the mature phage coat results in phage particles that present an antibody on their surface and contain the genetic material encoding the antibody. A phage library comprising such constructs is expressed in bacteria, and the library is screened for nGPCR-x-specific phage-antibodies using labeled or immobilized nGPCR-x as antigen-probe.

D. Human nGPCR-x-neutralizing antibodies from transgenic mice

Human nGPCR-x-neutralizing antibodies are generated in transgenic mice essentially as described in Bruggemann *et al.*, Immunol. Today 17(8):391-97 (1996) and Bruggemann *et al.*, Curr. Opin. Biotechnol. 8:455-58 (1997). Transgenic mice carrying human V-gene segments in germline configuration and that express these transgenes in their lymphoid tissue are immunized with a nGPCR-x composition using conventional immunization protocols. Hybridomas are generated using B cells from the immunized mice using conventional protocols and screened to identify hybridomas secreting anti-nGPCR-x human antibodies (*e.g.*, as described above).

EXAMPLE 10: ASSAYS TO IDENTIFY MODULATORS OF nGPCR-X ACTIVITY

Set forth below are several nonlimiting assays for identifying modulators (agonists and antagonists) of nGPCR-x activity. Among the modulators that can be

5

10

15

20

25

identified by these assays are natural ligand compounds of the receptor; synthetic analogs and derivatives of natural ligands; antibodies, antibody fragments, and/or antibody-like compounds derived from natural antibodies or from antibody-like combinatorial libraries; and/or synthetic compounds identified by high-throughput screening of libraries; and the like. All modulators that bind nGPCR-x are useful for identifying nGPCR-x in tissue samples (e.g., for diagnostic purposes, pathological purposes, and the like). Agonist and antagonist modulators are useful for upregulating and down-regulating nGPCR-x activity, respectively, to treat disease states characterized by abnormal levels of nGPCR-x activity. The assays may be performed using single putative modulators, and/or may be performed using a known agonist in combination with candidate antagonists (or visa versa).

A. cAMP Assays

In one type of assay, levels of cyclic adenosine monophosphate (cAMP) are measured in nGPCR-x-transfected cells that have been exposed to candidate modulator compounds. Protocols for cAMP assays have been described in the literature. (See, e.g., Sutherland et al., Circulation 37: 279 (1968); Frandsen et al., Life Sciences 18: 529-541 (1976); Dooley et al., Journal of Pharmacology and Experimental Therapeutics 283 (2): 735-41 (1997); and George et al., Journal of Biomolecular Screening 2 (4): 235-40 (1997)). An exemplary protocol for such an assay, using an Adenylyl Cyclase Activation FlashPlate® Assay from NEN™ Life Science Products, is set forth below.

Briefly, the nGPCR-x coding sequence (e.g., a cDNA or intronless genomic DNA) is subcloned into a commercial expression vector, such as pzeoSV2 (Invitrogen), and transiently transfected into Chinese Hamster Ovary (CHO) cells using known methods, such as the transfection protocol provided by Boehringer-Mannheim when supplying the FuGENE 6 transfection reagent. Transfected CHO cells are seeded into 96-well microplates from the FlashPlate® assay kit, which are coated with solid scintillant to which antisera to cAMP has been bound. For a control, some wells are seeded with wild type (untransfected) CHO cells. Other wells in the plate receive various amounts of a cAMP standard solution for use in creating a standard curve.

One or more test compounds (i.e., candidate modulators) are added to the cells in each well, with water and/or compound-free medium/diluent serving as a control or

5

10

15

20

25

controls. After treatment, cAMP is allowed to accumulate in the cells for exactly 15 minutes at room temperature. The assay is terminated by the addition of lysis buffer containing [125]-labeled cAMP, and the plate is counted using a Packard Topcount™ 96-well microplate scintillation counter. Unlabeled cAMP from the lysed cells (or from standards) and fixed amounts of [125]-cAMP compete for antibody bound to the plate. A standard curve is constructed, and cAMP values for the unknowns are obtained by interpolation. Changes in intracellular cAMP levels of cells in response to exposure to a test compound are indicative of nGPCR-x modulating activity. Modulators that act as agonists of receptors which couple to the G_s subtype of G proteins will stimulate production of cAMP, leading to a measurable 3-10 fold increase in cAMP levels. Agonists of receptors which couple to the G_{i/o} subtype of G proteins will inhibit forskolin-stimulated cAMP production, leading to a measurable decrease in cAMP levels of 50-100%. Modulators that act as inverse agonists will reverse these effects at receptors that are either constitutively active or activated by known agonists.

B. Aequorin Assays

In another assay, cells (e.g., CHO cells) are transiently co-transfected with both a nGPCR-x expression construct and a construct that encodes the photoprotein apoaquorin. In the presence of the cofactor coelenterazine, apoaquorin will emit a measurable luminescence that is proportional to the amount of intracellular (cytoplasmic) free calcium. (See generally, Cobbold, et al. "Aequorin measurements of cytoplasmic free calcium," In: McCormack J.G. and Cobbold P.H., eds., Cellular Calcium: A Practical Approach. Oxford:IRL Press (1991); Stables et al., Analytical Biochemistry 252: 115-26 (1997); and Haugland, Handbook of Fluorescent Probes and Research Chemicals. Sixth edition. Eugene OR: Molecular Probes (1996).)

In one exemplary assay, nGPCR-x is subcloned into the commercial expression vector pzeoSV2 (Invitrogen) and transiently co-transfected along with a construct that encodes the photoprotein apoaquorin (Molecular Probes, Eugene, OR) into CHO cells using the transfection reagent FuGENE 6 (Boehringer-Mannheim) and the transfection protocol provided in the product insert.

The cells are cultured for 24 hours at 37°C in MEM (Gibco/BRL, Gaithersburg, MD) supplemented with 10% fetal bovine serum, 2 mM glutamine, 10 U/ml penicillin and 10 µg/ml streptomycin, at which time the medium is changed to

5

10

15

20

25

serum-free MEM containing 5 μ M coelenterazine (Molecular Probes, Eugene, OR). Culturing is then continued for two additional hours at 37°C. Subsequently, cells are detached from the plate using VERSEN (Gibco/BRL), washed, and resuspended at 200,000 cells/ml in serum-free MEM.

Dilutions of candidate nGPCR-x modulator compounds are prepared in serum-free MEM and dispensed into wells of an opaque 96-well assay plate at 50 μ l/well. Plates are then loaded onto an MLX microtiter plate luminometer (Dynex Technologies, Inc., Chantilly, VA). The instrument is programmed to dispense 50 μ l cell suspensions into each well, one well at a time, and immediately read luminescence for 15 seconds. Dose-response curves for the candidate modulators are constructed using the area under the curve for each light signal peak. Data are analyzed with SlideWrite, using the equation for a one-site ligand, and EC50 values are obtained. Changes in luminescence caused by the compounds are considered indicative of modulatory activity. Modulators that act as agonists at receptors which couple to the G_q subtype of G proteins give an increase in luminescence of up to 100 fold. Modulators that act as inverse agonists will reverse this effect at receptors that are either constitutively active or activated by known agonists.

C. <u>Luciferase Reporter Gene Assay</u>

The photoprotein luciferase provides another useful tool for assaying for modulators of nGPCR-x activity. Cells (e.g., CHO cells or COS 7 cells) are transiently co-transfected with both a nGPCR-x expression construct (e.g., nGPCR-x in pzeoSV2) and a reporter construct which includes a gene for the luciferase protein downstream from a transcription factor binding site, such as the cAMP-response element (CRE), AP-1, or NF-kappa B. Agonist binding to receptors coupled to the G_s subtype of G proteins leads to increases in cAMP, thereby activating the CRE transcription factor and resulting in expression of the luciferase gene. Agonist binding to receptors coupled to the G_q subtype of G protein leads to production of diacylglycerol that activates protein kinase C, which activates the AP-1 or NF-kappa B transcription factors, in turn resulting in expression of the luciferase gene. Expression levels of luciferase reflect the activation status of the signaling events. (See generally, George et al., Journal of Biomolecular Screening 2(4): 235-240 (1997); and Stratowa et al., Current Opinion in Biotechnology 6: 574-581 (1995)).

5

10

15

20

25

Luciferase activity may be quantitatively measured using, e.g., luciferase assay reagents that are commercially available from Promega (Madison, WI).

5

10

15

20

25

30

BN\$00000 kW0 | 2136473A3 1 8

In one exemplary assay, CHO cells are plated in 24-well culture dishes at a density of 100,000 cells/well one day prior to transfection and cultured at 37°C in MEM (Gibco/BRL) supplemented with 10% fetal bovine serum, 2 mM glutamine, 10 U/ml penicillin and 10 µg/ml streptomycin. Cells are transiently co-transfected with both a nGPCR-x expression construct and a reporter construct containing the luciferase gene. The reporter plasmids CRE-luciferase, AP-1-luciferase and NFkappaB-luciferase may be purchased from Stratagene (LaJolla, CA). Transfections are performed using the FuGENE 6 transfection reagent (Boehringer-Mannheim) according to the supplier's instructions. Cells transfected with the reporter construct alone are used as a control. Twenty-four hours after transfection, cells are washed once with PBS pre-warmed to 37°C. Serum-free MEM is then added to the cells either alone (control) or with one or more candidate modulators and the cells are incubated at 37°C for five hours. Thereafter, cells are washed once with ice-cold PBS and lysed by the addition of 100 μ l of lysis buffer per well from the luciferase assay kit supplied by Promega. After incubation for 15 minutes at room temperature, 15 μ l of the lysate is mixed with 50 µl of substrate solution (Promega) in an opaque-white, 96-well plate, and the luminescence is read immediately on a Wallace model 1450 MicroBeta scintillation and luminescence counter (Wallace Instruments, Gaithersburg, MD).

Differences in luminescence in the presence versus the absence of a candidate modulator compound are indicative of modulatory activity. Receptors that are either constitutively active or activated by agonists typically give a 3 to 20-fold stimulation of luminescence compared to cells transfected with the reporter gene alone. Modulators that act as inverse agonists will reverse this effect.

D. Intracellular calcium measurement using FLIPR

Changes in intracellular calcium levels are another recognized indicator of G protein-coupled receptor activity, and such assays can be employed to screen for modulators of nGPCR-x activity. For example, CHO cells stably transfected with a nGPCR-x expression vector are plated at a density of 4 x 10⁴ cells/well in Packard black-walled, 96-well plates specially designed to discriminate fluorescence signals emanating from the various wells on the plate. The cells are incubated for 60 minutes

at 37°C in modified Dulbecco's PBS (D-PBS) containing 36 mg/L pyruvate and 1 g/L glucose with the addition of 1% fetal bovine serum and one of four calcium indicator dyes (Fluo-3TM AM, Fluo-4TM AM, Calcium GreenTM-1 AM, or Oregon GreenTM 488 BAPTA-1 AM), each at a concentration of 4 μ M. Plates are washed once with modified D-PBS without 1% fetal bovine serum and incubated for 10 minutes at 37°C to remove residual dye from the cellular membrane. In addition, a series of washes with modified D-PBS without 1% fetal bovine serum is performed immediately prior to activation of the calcium response.

A calcium response is initiated by the addition of one or more candidate receptor agonist compounds, calcium ionophore A23187 (10 μ M; positive control), or ATP (4 μ M; positive control). Fluorescence is measured by Molecular Device's FLIPR with an argon laser (excitation at 488 nm). (See, *e.g.*, Kuntzweiler *et al.*, Drug Development Research, 44(1):14-20 (1998)). The F-stop for the detector camera was set at 2.5 and the length of exposure was 0.4 milliseconds. Basal fluorescence of cells was measured for 20 seconds prior to addition of candidate agonist, ATP, or A23187, and the basal fluorescence level was subtracted from the response signal. The calcium signal is measured for approximately 200 seconds, taking readings every two seconds. Calcium ionophore A23187 and ATP increase the calcium signal 200% above baseline levels. In general, activated GPCRs increase the calcium signal approximately 10-15% above baseline signal.

E. Mitogenesis Assay

In a mitogenesis assay, the ability of candidate modulators to induce or inhibit nGPCR-x-mediated cell division is determined. (*See, e.g.*, Lajiness *et al.*, Journal of Pharmacology and Experimental Therapeutics 267(3): 1573-1581 (1993)). For example, CHO cells stably expressing nGPCR-x are seeded into 96-well plates at a density of 5000 cells/well and grown at 37°C in MEM with 10% fetal calf serum for 48 hours, at which time the cells are rinsed twice with serum-free MEM. After rinsing, $80~\mu l$ of fresh MEM, or MEM containing a known mitogen, is added along with $20~\mu l$ MEM containing varying concentrations of one or more candidate modulators or test compounds diluted in serum-free medium. As controls, some wells on each plate receive serum-free medium alone, and some receive medium containing 10% fetal bovine serum. Untransfected cells or cells transfected with vector alone also may serve as controls.

5

10

15

20

25

After culture for 16-18 hours, 1 μ Ci of [³H]-thymidine (2 Ci/mmol) is added to the wells and cells are incubated for an additional 2 hours at 37°C. The cells are trypsinized and collected on filter mats with a cell harvester (Tomtec); the filters are then counted in a Betaplate counter. The incorporation of [³H]-thymidine in serum-free test wells is compared to the results achieved in cells stimulated with serum (positive control). Use of multiple concentrations of test compounds permits creation and analysis of dose-response curves using the non-linear, least squares fit equation: $A = B \times [C/(D+C)] + G$ where A is the percent of serum stimulation; B is the maximal effect minus baseline; C is the EC₅₀; D is the concentration of the compound; and G is the maximal effect. Parameters B, C and G are determined by Simplex optimization.

Agonists that bind to the receptor are expected to increase [³H]-thymidine incorporation into cells, showing up to 80% of the response to serum. Antagonists that bind to the receptor will inhibit the stimulation seen with a known agonist by up to 100%.

F. [35S]GTPγS Binding Assay

Because G protein-coupled receptors signal through intracellular G proteins whose activity involves GTP binding and hydrolysis to yield bound GDP, measurement of binding of the non-hydrolyzable GTP analog [35 S]GTP γ S in the presence and absence of candidate modulators provides another assay for modulator activity. (See, e.g., Kowal et al., Neuropharmacology 37:179-187 (1998).)

In one exemplary assay, cells stably transfected with a nGPCR-x expression vector are grown in 10 cm tissue culture dishes to subconfluence, rinsed once with 5 ml of ice-cold $\text{Ca}^{2+}/\text{Mg}^{2+}$ -free phosphate-buffered saline, and scraped into 5 ml of the same buffer. Cells are pelleted by centrifugation (500 x g, 5 minutes), resuspended in TEE buffer (25 mM Tris, pH 7.5, 5 mM EDTA, 5 mM EGTA), and frozen in liquid nitrogen. After thawing, the cells are homogenized using a Dounce homogenizer (one ml TEE per plate of cells), and centrifuged at 1,000 x g for 5 minutes to remove nuclei and unbroken cells.

The homogenate supernatant is centrifuged at 20,000 x g for 20 minutes to isolate the membrane fraction, and the membrane pellet is washed once with TEE and resuspended in binding buffer (20 mM HEPES, pH 7.5, 150 mM NaCl, 10 mM

5

10

15

20

25

MgCl₂, 1 mM EDTA). The resuspended membranes can be frozen in liquid nitrogen and stored at -70°C until use.

Aliquots of cell membranes prepared as described above and stored at -70°C are thawed, homogenized, and diluted into buffer containing 20 mM HEPES, 10 mM MgCl₂, 1 mM EDTA, 120 mM NaCl, 10 μ M GDP, and 0.2 mM ascorbate, at a concentration of 10-50 μ g/ml. In a final volume of 90 μ l, homogenates are incubated with varying concentrations of candidate modulator compounds or 100 μ M GTP for 30 minutes at 30°C and then placed on ice. To each sample, 10 μ l guanosine 5'-O-(3[35 S]thio) triphosphate (NEN, 1200 Ci/mmol; [35 S]-GTP γ S), was added to a final concentration of 100-200 pM. Samples are incubated at 30°C for an additional 30 minutes, 1 ml of 10 mM HEPES, pH 7.4, 10 mM MgCl₂, at 4°C is added and the reaction is stopped by filtration.

Samples are filtered over Whatman GF/B filters and the filters are washed with 20 ml ice-cold 10 mM HEPES, pH 7.4, 10 mM MgCl₂. Filters are counted by liquid scintillation spectroscopy. Nonspecific binding of [35 S]-GTP γ S is measured in the presence of 100 μ M GTP and subtracted from the total. Compounds are selected that modulate the amount of [35 S]-GTP γ S binding in the cells, compared to untransfected control cells. Activation of receptors by agonists gives up to a five-fold increase in [35 S]GTP γ S binding. This response is blocked by antagonists.

G. MAP Kinase Activity Assay

Evaluation of MAP kinase activity in cells expressing a GPCR provides another assay to identify modulators of GPCR activity. (See, e.g., Lajiness et al., Journal of Pharmacology and Experimental Therapeutics 267(3):1573-1581 (1993) and Boulton et al., Cell 65:663-675 (1991).)

In one embodiment, CHO cells stably transfected with nGPCR-x are seeded into 6-well plates at a density of 70,000 cells/well 48 hours prior to the assay. During this 48-hour period, the cells are cultured at 37°C in MEM medium supplemented with 10% fetal bovine serum, 2 mM glutamine, 10 U/ml penicillin and 10 μ g/ml streptomycin. The cells are serum-starved for 1-2 hours prior to the addition of stimulants.

For the assay, the cells are treated with medium alone or medium containing either a candidate agonist or 200 nM Phorbol ester- myristoyl acetate (*i.e.*, PMA, a positive control), and the cells are incubated at 37°C for varying times. To stop the

10

15

20

25

reaction, the plates are placed on ice, the medium is aspirated, and the cells are rinsed with 1 ml of ice-cold PBS containing 1 mM EDTA. Thereafter, 200 μ l of cell lysis buffer (12.5 mM MOPS, pH 7.3, 12.5 mM glycerophosphate, 7.5 mM MgCl₂, 0.5 mM EGTA, 0.5 mM sodium vanadate, 1 mM benzamidine, 1 mM dithiothreitol, 10 μ g/ml leupeptin, 10 μ g/ml aprotinin, 2 μ g/ml pepstatin A, and 1 μ M okadaic acid) is added to the cells. The cells are scraped from the plates and homogenized by 10 passages through a 23 3/4 G needle, and the cytosol fraction is prepared by centrifugation at 20,000 x g for 15 minutes.

Aliquots (5-10 μ l containing 1-5 μ g protein) of cytosol are mixed with 1 mM MAPK Substrate Peptide (APRTPGGRR (SEQ ID NO: 129), Upstate Biotechnology, Inc., N.Y.) and 50 μ M [γ^{-32} P]ATP (NEN, 3000 Ci/mmol), diluted to a final specific activity of ~2000 cpm/pmol, in a total volume of 25 μ l. The samples are incubated for 5 minutes at 30°C, and reactions are stopped by spotting 20 μ l on 2 cm² squares of Whatman P81 phosphocellulose paper. The filter squares are washed in 4 changes of 1% H₃PO₄, and the squares are subjected to liquid scintillation spectroscopy to quantitate bound label. Equivalent cytosolic extracts are incubated without MAPK substrate peptide, and the bound label from these samples are subtracted from the matched samples with the substrate peptide. The cytosolic extract from each well is used as a separate point. Protein concentrations are determined by a dye binding protein assay (Bio-Rad Laboratories). Agonist activation of the receptor is expected to result in up to a five-fold increase in MAPK enzyme activity. This increase is blocked by antagonists.

H. [3H]Arachidonic Acid Release

The activation of GPCRs also has been observed to potentiate arachidonic acid release in cells, providing yet another useful assay for modulators of GPCR activity. (See, e.g., Kanterman et al., Molecular Pharmacology 39:364-369 (1991).) For example, CHO cells that are stably transfected with a nGPCR-x expression vector are plated in 24-well plates at a density of 15,000 cells/well and grown in MEM medium supplemented with 10% fetal bovine serum, 2 mM glutamine, 10 U/ml penicillin and $10 \mu g/ml$ streptomycin for 48 hours at 37°C before use. Cells of each well are labeled by incubation with [3 H]-arachidonic acid (Amersham Corp., 210 Ci/mmol) at 0.5 μ Ci/ml in 1 ml MEM supplemented with 10 mM HEPES, pH 7.5, and 0.5% fatty-

5

10

15

20

25

acid-free bovine serum albumin for 2 hours at 37°C. The cells are then washed twice with 1 ml of the same buffer.

Candidate modulator compounds are added in 1 ml of the same buffer, either alone or with 10 μ M ATP and the cells are incubated at 37°C for 30 minutes. Buffer alone and mock-transfected cells are used as controls. Samples (0.5 ml) from each well are counted by liquid scintillation spectroscopy. Agonists which activate the receptor will lead to potentiation of the ATP-stimulated release of [3 H]-arachidonic acid. This potentiation is blocked by antagonists.

I. Extracellular Acidification Rate

In yet another assay, the effects of candidate modulators of nGPCR-x activity are assayed by monitoring extracellular changes in pH induced by the test compounds. (See, e.g., Dunlop et al., Journal of Pharmacological and Toxicological Methods 40(1):47-55 (1998).) In one embodiment, CHO cells transfected with a nGPCR-x expression vector are seeded into 12 mm capsule cups (Molecular Devices Corp.) at 4×10^5 cells/cup in MEM supplemented with 10% fetal bovine serum, 2 mM L-glutamine, 10 U/ml penicillin, and 10 μ g/ml streptomycin. The cells are incubated in this medium at 37°C in 5% CO₂ for 24 hours.

Extracellular acidification rates are measured using a Cytosensor microphysiometer (Molecular Devices Corp.). The capsule cups are loaded into the sensor chambers of the microphysiometer and the chambers are perfused with running buffer (bicarbonate-free MEM supplemented with 4 mM L-glutamine, 10 units/ml penicillin, 10 μg/ml streptomycin, 26 mM NaCl) at a flow rate of 100 μl/minute. Candidate agonists or other agents are diluted into the running buffer and perfused through a second fluid path. During each 60-second pump cycle, the pump is run for 38 seconds and is off for the remaining 22 seconds. The pH of the running buffer in the sensor chamber is recorded during the cycle from 43-58 seconds, and the pump is re-started at 60 seconds to start the next cycle. The rate of acidification of the running buffer during the recording time is calculated by the Cytosoft program. Changes in the rate of acidification are calculated by subtracting the baseline value (the average of 4 rate measurements immediately before addition of a modulator candidate) from the highest rate measurement obtained after addition of a modulator candidate. The selected instrument detects 61 mV/pH unit. Modulators that act as agonists of the receptor result in an increase in the rate of extracellular acidification compared to the

5

10

15

20

25

rate in the absence of agonist. This response is blocked by modulators which act as antagonists of the receptor.

EXAMPLE 11: IN SITU HYBRIDIZATION

DNA Probe Preparation For nGPCR-11, -16, -40, -54, and -56

DNA probes for *in situ* hybridization were prepared as follows. Two sets of primer pairs were prepared. The first set has the sequence for T7 polymerase promoter on the 5' primer to make the sense RNA, and the second set has the T7 polymerase promoter sequence on the 3' primer to make the antisense RNA. PCR was performed in a 50 µl reaction containing 36.5 µl H₂O, 5µl 10xTT buffer (140 mM Ammonium Sulfate, 0.1 % gelatine, 0.6 M Tris-tricine pH 8.4), 5 µl 25mM MgCl₂, 2 µl 10 mM dNTP, 0.4 µl Incyte clone 1722192 DNA, 0.5 µl AmpliTaq (PE Applied Biosystems), and 0.3 µl oligo1 (1 mg/ml) and 0.3 µl oligo2 (1mg/ml)[to make the sense RNA], or 0.3 µl oligo3 (1 mg/ml) and 0.3 µl oligo4 (1mg/ml)[to make the antisense RNA]. The PCR reaction involved one cycle at 94°C for 2 min followed by 35 cycles at 94°C for 30 sec, 60°C for 30 sec, 72°C for 30 sec. The two PCR reactions were loaded onto a 1.2 % agarose gel. The DNA band was excised from the gel, placed in a GenElute Agarose spin column (Supelco) and spun for 10 min at maximum speed. The eluted DNA was EtOH precipitated and resuspended in transcription buffer. The primer sequences for each nGPCR tested are listed below.

For nGPCR-11, the sense primers were:

GCGTAATACGACTCACTATAGGGAGACCGCGTGTCTGCTAGACTCTATTTC C 3'(LW1658) (SEQ $\rm ID$ NO: 159), and:

5' TGCCACACTGATGCAACTCC 3' (LW1661) (SEQ ID NO: 160). The antisense primers were:

GCGTAATACGACTCACTATAGGGAGACCTGCCACACTGATGCAACTCC (LW1659) SEQ ID NO: 161) and.

5'GCGTGTCTGCTAGACTCTATTTCC 3' (LW1660) (SEQ ID NO: 162). The primer pairs yielded a product of 275bp.

For nGPCR-16, the sense primers were: 5'GCGTAATACGACTCACTATAGGGAGACCGCACGCCACTCTTTACTATCC C (LW1645) (SEQ ID NO: 163), and:

5

10

15

20

25

5' GCACAAAACACAATTCCATAAGCC 3' (LW1648) (SEQ ID NO: 164). The antisense primers were:

- 5'GCGTAATACGACTCACTATAGGGAGACCGCACAAAACACAATTCCATAA GCC 3' (LW1646) (SEQ ID NO: 165), and:
- 5 5' GCTACGCCACTCTTTACTATCCC 3'(LW1647) (SEQ ID NO: 166). The primer pairs yielded a product of 283 bp.

For nGPCR-40, the sense primers were:

- 5'GCGTAATACGACTCACTATAGGGAGACCTTATGAGCAGCAATTCATCCC 3'(LW1704) (SEQ ID NO: 167), and:
- 5'CACACCCACCAAGAAATCAG 3'(LW1707)(SEQ ID NO: 168). The antisense primers were:
 - 5'GCGTAATACGACTCACTATAGGGAGACCCACACCCACCAAGAAATCAG 3'(LW1705) (SEQ ID NO: 169), and:
- 5' TTATGAGCAGCAATTCATCCC 3' (LW1706) (SEQ ID NO: 170). The primer pairs yielded a product of 251bp.

For nGPCR-54, the sense primers were:

- 5'GCGTAATACGACTCACTATAGGGAGACCCGATTATCCACACTTTGACCC 3' (LW1803) (SEQ ID NO: 171), and:
- 5' CTGAAAGTTGTCGCTGACC 3' (LW1634) (SEQ ID NO: 172). The anti-sense primers were:
- GCGTAATACGACTCACTATAGGGAGACCCTGCTGAAAGTTGTCGCTGACC 3' (LW1804)(SEQ ID NO: 173), and:
- 5' CGATTATCCACACTTTGACCC 3' (LW1635) (SEQ ID NO: 174). The primer pairs yielded a product of 286 bp.
- For nGPCR-56, the sense primers were:
 - GCGTAATACGACTCACTATAGGGAGACCCTGTAAAATTCACACAAGCACC 3' (LW1763) (SEQ ID NO: 175), and:
 - 5'AGAAGACAGAGCAACCTCC 3' (LW1766) (SEQ ID NO: 176). The anti-sense primers were:
- 30 GCGTAATACGACTCACTATAGGGAGACCAGAAGACAGAGCAACCTCC (LW1764) (SEQ ID NO: 177) and:
 - CTGTAAAATTCACACAAGCACC (LW1765) (SEQ ID NO: 178). The primer pairs yielded a product of 272 bp.

DNA Probe Preparation For nGPCR-1

Probes for nGPCR-1 were prepared as above with the following modifications. Using a sense primer:

GCATGGATCCTCTTTGCTGTATTTCACCCTC) (LW1595) (SEQ ID NO: 179) and an antisense primer:

5'GCATGAATTCACAATGCCAGTGATAAGGAAG 3' (LW1596) (SEQ ID NO: 180), a 271 bp fragment was generated by PCR. The fragment was digested with *BamHI* and *EcoRI* and ligated into a BluescriptII vector that had been cut with *BamHI* and *EcoRI*. The orientation of the insert was such that T7 polymerase generates the anti-sense strand and T3 polymerase generates the sense strand.

Histochemistry

5

10

15

20

25

30

Coronal and sagittal oriented rat brain sections were cryosectioned (20 um thick) using a Reichert-Jung cryostat. The individual sections were thaw-mounted onto silanated, nuclease-free slides (CEL Associates, Inc., Houston, TX), and stored at -80°C. The sections were processed starting with post-fixation in cold 4% paraformaldehyde, rinsed in cold PBS, acetylated using acetic anhydride in triethanolamine buffer and dehydrated through 70%, 95%, and 100% alcohols at room temperature (RT). This was followed with delipidation in chloroform then rehydration in 100% and 95% alcohol at room temperature. Sections were air-dried prior to hybridization. Two PCR fragments (~250 bp) were generated, one that contained T7 polymerase on the 5' end (sense) and the other with T7 polymerase on the 3' end (antisense). The PCR fragments were labeled with 35S-UTP to yield a specific activity of 0.655 x 10⁶ cpm/pmol for antisense and 0.675 x 10⁶ cpm/pmol for sense probe. Both riboprobes were denatured and added to hybridization buffer containing 50% formamide, 10% dextran, 0.3M NaCl, 10 mM Tris, 1 mM EDTA, 1X Denhardts, and 10 mM DTT. Sequential brain cryosections were hybridized with 45 ul/slide of the sense and antisense riboprobe hybridization mixture, then covered with silanized glass coverslips. The sections were hybridized overnight (15-18 hrs) at 42°C in an incubator.

Coverslips were washed off the slides in 1X SSC, followed by RNase A treatment, and high temperature stringency washes (3X, 20 mins at 41°C) in 0.1X SSC. Slides were dehydrated with 70%, 95% NH₄OAc, and 100% NH₄OAc alcohols, air-dried and exposed to Kodak BioMax MR-1 film. After 9 days of exposure, the

film was developed. This was followed with coating selected tissue slides with Kodak NTB-2 nuclear track emulsion and storing the slides in the dark for 23 days. The slides were then developed and counterstained with hematoxylin. Emulsion-coated sections were analyzed microscopically to determine the specificity of labeling. Presence of autoradiographic grains (generated by antisense probe hybridization) over cell bodies (versus between cell bodies) was used as an index of specific hybridization.

Results

and red nucleus.

5

10

15

20

In situ hybridization results indicated localization in the following brain areas: nGPCR-1 was localized to the dentate gyrus of hippocampus, piriform cortex,

nGPCR-11 was localized to the piriform cortex, hippocampus, red nucleus, subthalamic nuclei, dorsal raphe, interpeduncular nucleus, and habenula. nGPCR-16 was localized to the cortex, piriform cortex, hippocampus, thalamus, subthalamic nuclei, hypothalamus, bed nucleus stria terminalis and posterior striatum. nGPCR-40 was localized to the cortex, piriform cortex, hippocampus, substantia nigra compacta, hypothalamus, laterial septus, bed nucleus stria terminalis, thalamus, ventral tegmental area, interpeduncular nucleus, dorsal raphe, medical geniculate, islands of Calleja, subthamalmic nuclei, choroid plexus. nGPCR-54 was localized to the piriform cortex and hippocampus, including the dentate gyrus, CA1 and CA3. nGPCR-56 was localized to the piriform cortex, cortex, interpeduncular nuceus, red nucleus, hippocampus, habenula, substantia nigra pars compacta, mamillary body stria terminalis,hypothalamus, subthamalmic nuclei, corsal raphe, and ventral tegmental area.

25

30

EXAMPLE 12: CHROMOSOMAL LOCALIZATION Methods

Chromosomal location of the genes encoding nGPCRs was determined using the Stanford G3 Radiation Hybrid Panel (Research Genetics, Inc., Huntsville, AL). This panel contains 83 radiation hybrid clones of the entire human genome created by the Stanford Human Genome Center. PCR reactions were assembled containing 25ng of DNA from each clone and the components of the Expand Hi-Fi PCR SystemTM (Roche Molecular Biochemicals, Indianapolis, IN) in a final reaction volume of 15 µl. PCR primers were synthesized by Genosys Corp., The Woodlands, TX. PCR

reactions were incubated in a GeneAmp 9700 PCR thermocycler (Perkin Elmer Applied Biosystems). The following cycling program was executed: Pre-soak at (94° for 3min.)(94° for 30 sec.)(52°C for 60 sec.)(72° for 2 min.)] for 35 cycles. PCR reaction products were then separated and analyzed by electrophoresis on a 2.0% agarose gel, and stained with ethidium bromide. Lanes were scored for the presence or absence of the expected PCR product and the results submitted to the Stanford Human Genome Center via e-mail for analysis (http://www-shgc.stanford.edu./RH/rhserverformnew.html).

nGPCR-40

5

10

15

20

25

30

PCR primers were designed based on the available sequence of the Celera sequence HUM_IDS|Contig|11000258115466. The forward primer used was:

5'ACAGCCCAAAGCCAAACAC3' (SEQ ID NO: 181). The reverse primer was:

5'CCGCAGGAGCAATG-AAAATCAG3' (SEQ ID NO: 182). This primer set will prime the synthesis of a 220 base pair fragment in the presence of the appropriate genomic DNA.

G3 Radiation Hybrid Panel Analysis places nGPCR-40 on chromosome 6, most nearly linked to Stanford marker SHGC-1836 with a LOD score of 11.84. This marker lies at position 6q21. In a genome scanning data set, Cao *et al.* (Genomics 1997 Jul 1: 43(1): 1-8) found excess allele sharing for markers on 6q13-q26. Greatest allele sharing was at interval 6q21-q22.3 with a maximum multipoint MLS value of 3.06 close to marker D6S278. Replication data from a second data set found maximum multipoint MLS at the interval D6S424-D6S275. These results provide suggestive evidence for a susceptibility locus for schizophrenia in chromosome 6q from two independent data sets.

nGPCR-54

PCR primers were designed based on the available sequence of the Celera sequence GA 11824020. The forward primer used was:

5'CTGTCTCTCTGTCCTCTTCC3', (SEQ ID NO: 183). The reverse primer used was:

5'GCACCGATCTTCATTGAATTTC3', (SEQ ID NO: 184). This primer set will prime the synthesis of a 145 base pair fragment in the presence of the appropriate genomic DNA.

G3 Radiation Hybrid Panel Analysis places nGPCR-54 on chromosome 13, most nearly linked to Stanford marker SHGC-68276 with a LOD score of 6.31. This marker lies at position 13q32. Numerous investigations have found significant suggestion of linkage of schizophrenia to this region of chromosome 13q32. See, for example, Brzustowicz *et al.*, Am J Hum Genet 1999 Oct; 65(4): 1096-1103; Blouin *et al.*, Nat Genet 1998 Sep; 20(1): 70-3; Shaw *et al.*, Am J Med Genet. 1998 Sep 7; 81(5): 364-76; Lin *et al.*, Hum Genet 1997 Mar; 99(3): 417-20; Pulver *et al.*, Cold Spring Harb Symp Quant Biol 1996; 61:797-814.

Genes localized to chromosomal regions in linkage with schizophrenia are candidate genes for disease susceptibility. Genes in these regions with the potential to play a biochemical/functional role in the disease process (like G protein coupled receptors) have a high probability of being a disease-modifying locus. nGPCR-40 and -54, because of their chromosomal location, are attractive targets therefore for screening ligands useful in modulating cellular processes involved in schizophrenia.

EXAMPLE 13: CLONE DEPOSIT INFORMATION

In accordance with the Budapest Treaty, clones of the present invention have been deposited at the Agricultural Research Culture Collection (NRRL) International Depository Authority, 1815 N. University Street, Peoria, Illinois 61604, U.S.A.

20 Accession numbers and deposit dates are provided below in Table 6.

5

10

Table 6: DEPOSIT INFORMATION

Clone	Accession Number NRRL	Budapest Treaty Deposit Date					
nGPCR-1 (SEQ ID NO:73)	B-30243	2000 Jan 18					
nGPCR -5 (SEQ ID NO: 75)	B-30244	2000 Jan 18					
nGPCR -16 (SEQ ID NO: 81)	B-30245	2000 Jan 18					
nGPCR -11 (SEQ ID NO: 79)	B-30258	2000 Feb 02					
nGPCR -17 (SEQ ID NO: 23)	B-30259	2000 Feb 03					
nGPCR -9 (SEQ ID NO: 77)	B-30262	2000 Feb 22					
nGPCR -58 (SEQ ID NO: 91)	B-30274	2000 March 23					
nGPCR -56 (SEQ ID NO: 89)	B-30288	2000 May 5					
nGPCR -3 (SEQ ID NO:185)	B-30290	2000 May 5					
nGPCR -54 (SEQ ID NO: 85)	B-30291	2000 May 5					
nGPCR -40 (SEQ ID NO: 83*)	B-30299N	2000 June 02					

^{*} The clone deposited with NRLL Accession Number B30299N comprises a sequence identical to SEQ ID NO:83 but with the substitution of an "A" at nucleotide position 10.

Example 14 - Using nGPCR-x proteins to isolate neurotransmitters

The isolated nGPCR-x proteins, particularly nGPCR-1, nGPCR-3, nGPCR-9, nGPCR-11, nGPCR-16, nGPCR-40, nGPCR-54, nGPCR-56, and nGPCR-58, (SEQ ID NOS: SEQ ID NO: 2, SEQ ID NO: 74; SEQ ID NO: 4, SEQ ID NO: 186; SEQ ID NO:10, SEQ ID NO:78; SEQ ID NO:12, SEQ ID NO:80; SEQ ID NO: 22, SEQ ID NO:82, SEQ ID NO:54, SEQ ID NO:84; SEQ ID NO:60, SEQ ID NO: 86; SEQ ID NO:64, SEQ ID NO: 88, SEQ ID NO:90; SEQ ID NO:68, SEQ ID NO: 92, and SEQ ID NO:94, respectively) can be used to isolate novel or known neurotransmitters (Saito et al., Nature 400: 265-269, 1999). The cDNAs that encode the isolated nGPCR-x can be cloned into mammalian expression vectors and used to stably or transiently transfect mammalian cells including CHO, Cos or HEK293 cells. Receptor expression can be determined by Northern blot analysis of transfected cells and identification of an appropriately sized mRNA band (predicted size from the cDNA). Brain regions shown by mRNA analysis to express each of the nGPCR-x

proteins could be processed for peptide extraction using any of several protocols

10

15

((Reinsheidk R.K. et al., Science 270: 243-247, 1996; Sakurai, T., et al., Cell 92; 573-585, 1998; Hinuma, S., et al., Nature 393: 272-276, 1998). Chromotographic fractions of brain extracts could be tested for ability to activate nGPCR-x proteins by measuring second messenger production such as changes in cAMP production in the presence or absence of forskolin, changes in inositol 3-phosphate levels, changes in intracellular calcium levels or by indirect measures of receptor activation including receptor stimulated mitogenesis, receptor mediated changes in extracellular acidification or receptor mediated changes in reporter gene activation in response to cAMP or calcium (these methods should all be referenced in other sections of the patent). Receptor activation could also be monitored by co-transfecting cells with a chimeric GI_{q/i3} to force receptor coupling to a calcium stimulating pathway (Conklin et al., Nature 363; 274-276, 1993). Neurotransmitter mediated activation of receptors could also be monitored by measuring changes in [35 S]-GTPKS binding in membrane fractions prepared from transfected mammalian cells. This assay could also be performed using baculoviruses containing nGPCR-x proteins infected into SF9 insect cells.

The neurotransmitter which activates nGPCR-x proteins can be purified to homogeneity through successive rounds of purification using nGPCR-x proteins activation as a measurement of neurotransmitter activity. The composition of the neurotransmitter can be determined by mass spectrometry and Edman degradation if peptidergic. Neurotransmitters isolated in this manner will be bioactive materials which will alter neurotransmission in the central nervous system and will produce behavioral and biochemical changes.

Example 15 - Using nGPCR-x proteins to isolate and purify G proteins

cDNAs encoding nGPCR-x proteins are epitope-tagged at the amino terminuus end of the cDNA with the cleavable influenza-hemagglutinin signal sequence followed by the FLAG epitope (IBI, New Haven, CT). Additionally, these sequences are tagged at the carboxyl terminus with DNA encoding six histidine residues. (Amino and Carboxyl Terminal Modifications to Facilitate the Production and Purification of a G Protein-Coupled Receptor, B.K. Kobilka, *Analytical Biochemistry*, Vol. 231, No. 1, Oct 1995, pp. 269-271). The resulting sequences are cloned into a baculovirus expression vector such as pVL1392 (Invitrogen). The baculovirus expression vectors are used to infect SF-9 insect cells as described (Guan,

35

5

10

15

20

25

X. M., Kobilka, T. S., and Kobilka, B. K. (1992) *J. Biol. Chem.* **267**, 21995-21998). Infected SF-9 cells could be grown in 1000-ml cultures in SF900 II medium (Life Technologies, Inc.) containing 5% fetal calf serum (Gemini, Calabasas, CA) and 0.1 mg/ml gentamicin (Life Technologies, Inc.) for 48 hours at which time the cells could be harvested. Cell membrane preparations could be separated from soluble proteins following cell lysis. nGPCR-x protein purification is carried out as described for purification of the 92 receptor (Kobilka, Anal. Biochem., 231 (1): 269-271, 1995) including solubilization of the membranes in 0.8-1.0 % *n*-dodecyl -D-maltoside (DM) (CalBiochem, La Jolla, CA) in buffer containing protease inhibitors followed by Nicolumn chromatography using chelating SepharoseTM (Pharmacia, Uppsala, Sweden). The eluate from the Ni-column is further purified on an M1 anti-FLAG antibody column (IBI). Receptor containing fractions are monitored by using receptor specific antibodies following western blot analysis or by SDS-PAGE analysis to look for an appropriate sized protein band (appropriate size would be the predicted molecular weight of the protein).

5

10

15

20

25

SNSDDO T KWO

This method of purifying G protein is particularly useful to isolate G proteins that bind to the nGPCR-x proteins in the absence of an activating ligand.

Some of the preferred embodiments of the invention described above are outlined below and include, but are not limited to, the following embodiments. As those skilled in the art will appreciate, numerous changes and modifications may be made to the preferred embodiments of the invention without departing from the spirit of the invention. It is intended that all such variations fall within the scope of the invention.

The entire disclosure of each publication cited herein is hereby incorporated by reference.

What is claimed is:

5

1. An isolated nucleic acid molecule comprising a nucleotide sequence that encodes a polypeptide comprising an amino acid sequence homologous to even numbered sequences selected from the group consisting of: SEQ ID NO:2 to SEQ ID NO:94, SEQ ID NO:186, and fragments thereof; said nucleic acid molecule encoding at least a portion of nGPCR-x.

- 10 2. The isolated nucleic acid molecule of claim 1 comprising a sequence that encodes a polypeptide comprising even numbered sequences selected from the group consisting of SEQ ID NO:2 to SEQ ID NO:94, SEQ ID NO:186, and fragments thereof.
- 15 3. The isolated nucleic acid molecule of claim 1 comprising a sequence homologous to odd numbered sequences selected from the group consisting of SEQ ID NO:1 to SEQ ID NO:93, SEQ ID NO:185 and fragments thereof.
- The isolated nucleic acid molecule of claim 1 comprising a sequence selected
 from the group of odd numbered sequences consisting of SEQ ID NO:1 to SEQ ID
 NO: 93, SEQ ID NO:185 and fragments thereof.
 - 5. The isolated nucleic acid molecule of claim 4 comprising a sequence selected from the group of odd numbered sequences consisting of SEQ ID NO:1 to SEQ ID NO:93 and SEQ ID NO:185.
 - 6. The isolated nucleic acid molecule of claim 4 wherein said nucleotide sequence is selected from the group consisting of: SEQ ID NO:1, SEQ ID NO:73, SEQ ID NO:9, SEQ ID NO:77, SEQ ID NO:11, SEQ ID NO:79, SEQ ID NO:21, SEQ ID NO:81 SEQ ID NO:53, SEQ ID NO:83, SEQ ID NO:59, SEQ ID NO:85, SEQ ID NO:63, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:67, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:3, and SEQ ID NO:185.
- 7. The isolated nucleic acid molecule of claim 4 wherein said nucleotide sequence is selected from the group consisting of: SEQ ID NO:73, SEQ ID NO:77,

25

SEQ ID NO:79, SEQ ID NO:81 SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:89, SEO ID NO:93 and SEQ ID NO:185.

- 8. The isolated nucleic acid molecule of claim 1 wherein said nucleic acid molecule is DNA.
 - 9. The isolated nucleic acid molecule of claim 1 wherein said nucleic acid molecule is RNA.
- 10. An expression vector comprising a nucleic acid molecule of any one of claims 1 to 5.
 - 11. The expression vector of claim 10 wherein said nucleic acid molecule comprises a sequence selected from the group of odd numbered sequences consisting of SEQ ID NO:1 to SEQ ID NO:93 and SEQ ID NO:185.
 - 12. The expression vector of claim 10 wherein said nucleic acid molecule comprises a nucleotide sequence selected from the group consisting of: SEQ ID NO:1, SEQ ID NO:73, SEQ ID NO:9, SEQ ID NO:77, SEQ ID NO:11, SEQ ID NO:79, SEQ ID NO: 21, SEQ ID NO:81 SEQ ID NO:53, SEQ ID NO:83, SEQ ID NO:59, SEQ ID NO:85, SEQ ID NO:63, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:67, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO: 3, and SEQ ID NO: 185.
- The expression vector of claim 10 wherein said nucleotide sequence is
 selected from the group consisting of: SEQ ID NO: 73, SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81 SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:89, SEQ ID NO:93 and SEQ ID NO: 185.
 - 14. The expression vector of claim 10 wherein said vector is a plasmid.
 - 15. The expression vector of claim 10 wherein said vector is a viral particle.
 - 16. The expression vector of claim 15 wherein said vector is selected from the group consisting of adenoviruses, baculoviruses, parvoviruses, herpesviruses,

30

15

poxviruses, adeno-associated viruses, Semliki Forest viruses, vaccinia viruses, and retroviruses.

17. The expression vector of claim 10 wherein said nucleic acid molecule is
5 operably connected to a promoter selected from the group consisting of simian virus
40, mouse mammary tumor virus, long terminal repeat of human immunodeficiency
virus, maloney virus, cytomegalovirus immediate early promoter, Epstein Barr virus,
rous sarcoma virus, human actin, human myosin, human hemoglobin, human muscle
creatine, and human metalothionein.

18. A host cell transformed with an expression vector of claim 10.

- 19. The transformed host cell of claim 18 wherein said cell is a bacterial cell.
- 15 20. The transformed host cell of claim 19 wherein said bacterial cell is E. coli.
 - 21. The transformed host cell of claim 18 wherein said cell is yeast.
 - 22. The transformed host cell of claim 21 wherein said yeast is S. cerevisiae.
 - 23. The transformed host cell of claim 18 wherein said cell is an insect cell.
 - 24. The transformed host cell of claim 23 wherein said insect cell is S. frugiperda.
- 25. The transformed host cell of claim 18 wherein said cell is a mammalian cell.
 - 26. The transformed host cell of claim 25 wherein mammalian cell is selected from the group consisting of chinese hamster ovary cells, HeLa cells, African green monkey kidney cells, human 293 cells, and murine 3T3 fibroblasts.
 - 27. An isolated nucleic acid molecule comprising a nucleotide sequence complementary to at least a portion of a sequence selected from the group of odd numbered sequences consisting of SEQ ID NO:1 to SEQ ID NO:93 and SEQ ID NO:185, said portion comprising at least 10 nucleotides.

30

10

28. The nucleic acid molecule of claim 27 wherein said molecule is an antisense oligonucleotide directed to a region of a sequence selected from the group of odd numbered sequences consisting of SEQ ID NO:1 to SEQ ID NO:93 and SEQ ID NO:185.

- 29. The nucleic acid molecule of claim 28 wherein said oligonucleotide is directed to a regulatory region of a sequence selected from the group of odd numbered sequences consisting of SEQ ID NO:1 to SEQ ID NO:93 and SEQ ID NO:185.
- 30. The nucleic acid molecule of claim 27 wherein said molecule is an antisense oligonucleotide directed to a region of nucleotide sequence selected from the group consisting of: SEQ ID NO: 73, SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81 SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:89, SEQ ID NO:93 and SEQ ID NO: 185.
- 31. A composition comprising a nucleic acid molecule of any one of claims 1 to 5 or 27 and an acceptable carrier or diluent.
- 32. A composition comprising a recombinant expression vector of claim 10 and an acceptable carrier or diluent.
 - 33. A method of producing a polypeptide that comprises a sequence selected from the group of even numbered sequences consisting SEQ ID NO: 2 to SEQ ID NO: 94 and SEQ ID NO: 186, and homologs and fragments thereof, said method comprising the steps of:
 - a) introducing a recombinant expression vector of claim 10 into a compatible host cell;
 - b) growing said host cell under conditions for expression of said polypeptide; and
 - c) recovering said polypeptide.
 - 34. The method of claim 33 wherein said host cell is lysed and said polypeptide is recovered from the lysate of said host cell.

5

10

15

25

35. The method of claim 33 wherein said polypeptide is recovered by purifying the culture medium without lysing said host cell.

- 36. An isolated polypeptide encoded by a nucleic acid molecule of claim 1.
- 37. The polypeptide of claim 36 wherein said polypeptide comprises a sequence selected from the group of even numbered sequences consisting SEQ ID NO:2 to SEQ ID NO:94 and SEQ ID NO:186.
- 10 38. The polypeptide of claim 36 wherein said polypeptide comprises an amino acid sequence homologous to a sequence selected from the group of even numbered sequences consisting of SEQ ID NO:2 to SEQ ID NO:94 and SEQ ID NO:186.
- 39. The polypeptide of claim 36 wherein said sequence homologous to a sequence selected from the group of even numbered sequences consisting of SEQ ID NO:2 to SEQ ID NO:94 and SEQ ID NO:186 comprises at least one conservative amino acid substitution compared to the even numbered sequences in the group of even numbered sequences consisting of SEQ ID NO: 2 to SEQ ID NO: 94 and SEQ ID NO: 186.
- 20 40. The polypeptide of claim 36 wherein said polypeptide comprises a fragment of a polypeptide with a sequence selected from the group of even numbered sequences consisting of SEO ID NO:2 to SEQ ID NO:94 and SEQ ID NO:186.
- 41. The polypeptide of claim 36 wherein said polypeptide comprises an amino acid sequence selected from the group consisting of: SEQ ID NO: 2, SEQ ID NO: 74; SEQ ID NO: 4, SEQ ID NO: 186; SEQ ID NO:10, SEQ ID NO:78; SEQ ID NO:12, SEQ ID NO:80; SEQ ID NO: 22, SEQ ID NO:82; SEQ ID NO:54, SEQ ID NO:84; SEQ ID NO:60, SEQ ID NO: 86; SEQ ID NO:64, SEQ ID NO: 88, SEQ ID NO:90; SEQ ID NO:68, SEQ ID NO: 92, and SEQ ID NO:94.
 - 42. The polypeptide of claim 36 wherein said polypeptide comprises an amino acid sequence selected from the group consisting of: SEQ ID NO: 74; SEQ ID NO: 186; SEQ ID NO:78; SEQ ID NO:80; SEQ ID NO:82; SEQ ID NO:84; SEQ ID NO:86; SEQ ID NO:90; and SEQ ID NO:94.

30

43. A composition comprising a polypeptide of claim 36 and an acceptable carrier or diluent.

- 5 44. An isolated antibody which binds to an epitope on a polypeptide of claim 36.
 - 45. The antibody of claim 44 wherein said antibody is a monoclonal antibody.
 - 46. A composition comprising an antibody of claim 44 and an acceptable carrier or diluent.
 - 47. A method of inducing an immune response in a mammal against a polypeptide of claim 36 comprising administering to said mammal an amount of said polypeptide sufficient to induce said immune response.
 - 48. A method for identifying a compound which binds nGPCR-x comprising the steps of:
 - a) contacting nGPCR-x with a compound; and
 - b) determining whether said compound binds nGPCR-x.
 - The method of claim 48 wherein the nGPCR-x comprises an amino acid sequence selected from the group consisting of: SEQ ID NO: 2, SEQ ID NO: 74; SEQ ID NO: 4, SEQ ID NO: 186; SEQ ID NO:10, SEQ ID NO:78; SEQ ID NO:12, SEQ ID NO:80; SEQ ID NO: 22, SEQ ID NO:82; SEQ ID NO:54, SEQ ID NO:84; SEQ ID NO:60, SEQ ID NO: 86; SEQ ID NO:64, SEQ ID NO: 88, SEQ ID NO:90; SEQ ID NO:68, SEQ ID NO: 92, and SEQ ID NO:94.
 - 50. The method of claim 48 wherein the nGPCR-x comprises an amino acid sequence selected from the group consisting of: SEQ ID NO: 74; SEQ ID NO: 186; SEQ ID NO:78; SEQ ID NO:80; SEQ ID NO:82; SEQ ID NO:84; SEQ ID NO: 86; SEQ ID NO:90; and SEQ ID NO:94.
 - 51. The method of claim 48 wherein binding of said compound to nGPCR-x is determined by a protein binding assay.

10

15

20

25

52. The method of claim 48 wherein said protein binding assay is selected from the group consisting of a gel-shift assay, Western blot, radiolabeled competition assay, phage-based expression cloning, co-fractionation by chromatography, co-precipitation, cross linking, interaction trap/two-hybrid analysis, southwestern analysis, and ELISA.

- 53. A compound identified by the method of claim 48.
- 10 54. A method for identifying a compound which binds a nucleic acid molecule encoding nGPCR-x comprising the steps of:
 - a) contacting said nucleic acid molecule encoding nGPCR-x with a compound; and
- b) determining whether said compound binds said nucleic acid
 molecule.
 - 55. The method of claim 54 wherein binding is determined by a gel-shift assay.
 - 56. A compound identified by the method of claim 54.
 - 57. A method for identifying a compound which modulates the activity of nGPCR-x comprising the steps of:
 - a) contacting nGPCR-x with a compound; and
 - b) determining whether nGPCR-x activity has been modulated.
 - 58. The method of claim 57 wherein the nGPCR-x comprises an amino acid sequence selected from the group consisting of: SEQ ID NO: 2, SEQ ID NO: 74; SEQ ID NO: 4, SEQ ID NO: 186; SEQ ID NO:10, SEQ ID NO:78; SEQ ID NO:12, SEQ ID NO:80; SEQ ID NO: 22, SEQ ID NO:82; SEQ ID NO:54, SEQ ID NO:84; SEQ ID NO:60, SEQ ID NO: 86; SEQ ID NO:64, SEQ ID NO: 88, SEQ ID NO:90; SEQ ID NO:68, SEQ ID NO: 92, and SEQ ID NO:94.
 - 59. The method of claim 57 wherein the nGPCR-x comprises an amino acid sequence selected from the group consisting of: SEQ ID NO: 74; SEQ ID NO: 186;

5

20

25

SEQ ID NO:78; SEQ ID NO:80; SEQ ID NO:82; SEQ ID NO:84; SEQ ID NO: 86; SEQ ID NO:90; and SEQ ID NO:94.

- 60. The method of claim 57 wherein said activity is neuropeptide binding.
- 61. The method of claim 57 wherein said activity is neuropeptide signaling.
- 62. A compound identified by the method of claim 57.
- 10 63. A method of identifying an animal homolog of nGPCR-x comprising the steps:
 - a) comparing the nucleic acid sequences of the animal with a sequence selected from the group of odd numbered sequence consisting of SEQ ID NO: 1 to SEQ ID NO: 93, SEQ ID NO: 185, and portions thereof, said portions being at least 10 nucleotides; and
 - b) identifying nucleic acid sequences of the animal that are homologous to said sequence selected from the group of odd numbered sequence consisting of SEQ ID NO: 1 to SEQ ID NO: 93, SEQ ID NO: 185, and portions thereof.
 - 64. The method of claim 63 wherein comparing the nucleic acid sequences of the animal with a sequence selected from the group of odd numbered sequence consisting of SEQ ID NO: 1 to SEQ ID NO: 93, SEQ ID NO: 185, and portions thereof, said portions being at least 10 nucleotides is performed by DNA hybridization.
 - 65. The method of claim 63 wherein comparing the nucleic acid sequences of the animal with a sequence selected from the group of odd numbered sequence consisting of SEQ ID NO: 1 to SEQ ID NO: 93, SEQ ID NO: 185, and portions thereof, said portions being at least 10 nucleotides is performed by computer homology search.
 - 66. A method of screening a human subject to diagnose a disorder affecting the brain or genetic predisposition therefor, comprising the steps of:
 - (a) assaying nucleic acid of a human subject to determine a presence or an absence of a mutation altering an amino acid sequence, expression, or biological

5

15

20

25

activity of at least one nGPCR that is expressed in the brain, wherein the nGPCR comprises an amino acid sequence selected from the group consisting of: SEQ ID NO:74, SEQ ID NO:186, SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86, SEQ ID NO:90, and SEQ ID NO:94, and allelic variants thereof, and wherein the nucleic acid corresponds to a gene encoding the nGPCR; and

(b) diagnosing the disorder or predisposition from the presence or absence of said mutation, wherein the presence of a mutation altering the amino acid sequence, expression, or biological activity of the nGPCR in the nucleic acid correlates with an increased risk of developing the disorder.

67. A method according to claim 66, wherein the nGPCR is nGPCR-40 comprising an amino acid sequence set forth in SEQ ID NO:84 or an allelic variant thereof.

- 15 68. A method according to claim 66, wherein the nGPCR is nGPCR-54 comprising an amino acid sequence set forth in SEQ ID NO:86 or an allelic variant thereof.
 - 69. A method according to claim 66, wherein the disease is schizophrenia.

70. A method according to claim 66, wherein the assaying step comprises at least one procedure selected from the group consisting of:

a) comparing nucleotide sequences from the human subject and reference sequences and determining a difference of either

at least a nucleotide of at least one codon between the nucleotide sequences from the human subject that encodes an nGPCR-40 allele and an nGPCR-40 reference sequence, or

at least a nucleotide of at least one codon between the nucleotide sequences from the human subject that encodes an nGPCR-54 allele and an nGPCR-54 reference sequence;

(b) performing a hybridization assay to determine whether nucleic acid from the human subject has a nucleotide sequence identical to or different from one or more reference sequences;

5

10

20

25

(c) performing a polynucleotide migration assay to determine whether nucleic acid from the human subject has a nucleotide sequence identical to or different from one or more reference sequences; and

- (d) performing a restriction endonuclease digestion to determine whether nucleic acid from the human subject has a nucleotide sequence identical to or different from one or more reference sequences.
 - 71. A method according to claim 70 wherein the assaying step comprises: performing a polymerase chain reaction assay to amplify nucleic acid comprising nGPCR-40 or nGPCR-54 coding sequence, and determining nucleotide sequence of the amplified nucleic acid.
 - 72. A method of screening for an nGPCR-40 or nGPCR-54 hereditary schizophrenia genotype in a human patient, comprising the steps of:
- (a) providing a biological sample comprising nucleic acid from said patient, said nucleic acid including sequences corresponding to allelles of nGPCR-40 or nGPCR-54; and
 - (b) detecting the presence of one or more mutations in the nGPCR-40 allelle or the nGPCR-54 allelle;

wherein the presence of a mutation in an nGPCR-40 allelle or nGPCR-54 allele is indicative of a hereditary schizophrenia genotype.

- 73. The method according to claim 72 wherein said biological sample is a cell sample.
- 74. The method according to claim 72 wherein said detecting the presence of a mutation comprises sequencing at least a portion of said nucleic acid, said portion comprising at least one codon of said nGPCR-40 or nGPCR-54 alleles.
- 30 75. The method according to claim 72 wherein said nucleic acid is DNA.
 - 76. The method according to claim 72 wherein said nucleic acid is RNA.

5

10

15

20

77. A kit for screening a human subject to diagnose schizophrenia or a genetic predisposition therefor, comprising, in association:

- (a) an oligonucleotide useful as a probe for identifying polymorphisms in a human nGPCR-40 gene or a human nGPCR-54 gene, the oligonucleotide comprising 6-50 nucleotides in a sequence that is identical or complementary to a sequence of a wild type human nGPCR-40 or nGPCR-54 gene sequence or nGPCR-40 or nGPCR-54 coding sequence, except for one sequence difference selected from the group consisting of a nucleotide addition, a nucleotide deletion, or nucleotide substitution; and
- (b) a media packaged with the oligonucleotide, said media containing information for identifying polymorphisms that correlate with schizophrenia or a genetic predisposition therefor, the polymorphisms being identifiable using the oligonucleotide as a probe.
- 15 78. A method of identifying a nGPCR allelic variant that correlates with a mental disorder, comprising steps of:
 - (a) providing a biological sample comprising nucleic acid from a human patient diagnosed with a mental disorder, or from the patient's genetic progenitors or progeny;
 - (b) detecting in the nucleic acid the presence of one or more mutations in an nGPCR that is expressed in the brain, wherein the nGPCR comprises an amino acid sequence selected from the group consisting of SEQ ID NO:74, SEQ ID NO:186, SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86, SEQ ID NO:90, and SEQ ID NO:94, and allelic variants thereof, and wherein the nucleic acid includes sequence corresponding to the gene or genes encoding nGPCR;

wherein the one or more mutations detected indicates an allelic variant that correlates with a mental disorder.

79. A method according to claim 78, wherein the disorder is schizophrenia, and wherein the at least one nGPCR is nGPCR-40, nGPCR-54, or an allelic variant thereof.

10

20

25

80. A purified and isolated polynucleotide comprising a nucleotide sequence encoding an nGPCR-40 or nGPCR-54 allelic variant identified according to claim 79.

- 81. A host cell transformed or transfected with a polynucleotide according to claim 80 or with a vector comprising the polynucleotide.
- 82. A purified polynucleotide comprising a nucleotide sequence encoding nGPCR-40 or nGPCR-54 of a human with schizophrenia;

wherein said polynucleotide hybridizes to the complement of SEQ ID NO:83 or of SEQ ID NO:85 under the following hybridization conditions:

- (a) hybridization for 16 hours at 42 °C in a hybridization solution comprising 50% formamide, 1% SDS, 1 M NaC1, 10% dextran sulfate and
- (b) washing 2 times for 30 minutes at 60°C in a wash solution comprising 0.1x SSC and 1% SDS; and

wherein the polynucleotide that encodes nGPCR-40 or nGPCR-54 amino acid sequence of the human differs from SEQ ID NO:84 or SEQ ID NO:86 by at least one residue.

- 83. A vector comprising a polynucleotide according to claim 82.
- 84. A host cell that has been transformed or transfected with a polynucleotide according to claim 82 and that expresses the nGPCR-40 or nGPCR-54 protein encoded by the polynucleotide.
- 25 85. A host cell according to claim 84 that has been co-transfected with a polynucleotide encoding the nGPCR-40 or nGPCR-54 amino acid sequence set forth in SEQ ID NO:84 or SEQ ID NO:86 and that expresses the nGPCR-40 or nGPCR-54 protein having the amino acid sequence set forth in SEQ ID NO:84 or SEQ ID NO:86.
- 86. A method for identifying a modulator of biological activity of nGPCR-40 or nGPCR-54 comprising the steps of:
 - a) contacting a cell according to claim 84 in the presence and in the absence of a putative modulator compound;

5

10

15

b) measuring nGPCR-40 or nGPCR-54 biological activity in the cell;

wherein decreased or increased nGPCR-40 or nGPCR-54 biological activity in the presence versus absence of the putative modulator is indicative of a modulator of biological activity.

- 87. A method to identify compounds useful for the treatment of schizophrenia, said method comprising steps of:
- (a) contacting a composition comprising nGPCR-40 with a compound suspected of binding nGPCR-40 or contacting a composition comprising nGPCR-54 with a compound suspected of binding nGPCR-54;
- (b) detecting binding between nGPCR-40 and the compound suspected of binding nGPCR-40 or between nGPCR-54 and the compound suspected of binding nGPCR-54;

wherein compounds identified as binding nGPCR-40 or nGPCR-54 are candidate compounds useful for the treatment of schizophrenia.

- 88. A method for identifying a compound useful as a modulator of binding between nGPCR-40 and a binding partner of nGPCR-40 or between nGPCR-54 and a binding partner of nGPCR-54 comprising the steps of:
- (a) contacting the binding partner and a composition comprising nGPCR-40 or nGPCR-54 in the presence and in the absence of a putative modulator compound;
- (b) detecting binding between the binding partner and nGPCR-40 or nGPCR-54;

wherein decreased or increased binding between the binding partner and nGPCR-40 or nGPCR-54 in the presence of the putative modulator, as compared to binding in the absence of the putative modulator is indicative a modulator compound useful for the treatment of schizophrenia.

89. A method according to claim 87 or 88 wherein the composition comprises a cell expressing nGPCR-40 or nGPCR-54 on its surface.

5

10

15

20

25

90. An method according to claim 89 wherein the composition comprises a cell transformed or transfected with a polynucleotide that encodes nGPCR-40 or nGPCR-54.

- 5 91. A method of purifying a G protein from a sample containing said G protein comprising the steps of:
 - a) contacting said sample with a polypeptide of claim 1 for a time sufficient to allow said G protein to form a complex with said polypeptide;
- b) isolating said complex from remaining components of said sample;
 - c) maintaining said complex under conditions which result in dissociation of said G protein from said polypeptide; and
 - d) isolating said G protein from said polypeptide.
- 15 92. The method of claim 91 wherein said sample comprises an amino acid sequence selected from the group of even numbered sequences consisting of SEQ ID NO:2 to SEQ ID NO:94 and SEQ ID NO:186.
 - 93. The method of claim 91 wherein said polypeptide comprises an amino acid sequence homologous to a sequence selected from the group of even numbered sequences consisting of SEQ ID NO:2 to SEQ ID NO:94 and SEQ ID NO:186.
 - 94. The method of claim 91 wherein said polypeptide comprises an amino acid sequence selected from the group consisting of: SEQ ID NO: 2, SEQ ID NO: 74; SEQ ID NO: 4, SEQ ID NO: 186; SEQ ID NO:10, SEQ ID NO:78; SEQ ID NO:12, SEQ ID NO:80; SEQ ID NO: 22, SEQ ID NO:82; SEQ ID NO:54, SEQ ID NO:84; SEQ ID NO:60, SEQ ID NO: 86; SEQ ID NO:64, SEQ ID NO: 88, SEQ ID NO:90; SEQ ID NO:68, SEQ ID NO: 92, and SEQ ID NO:94.
- 30 95. The method of claim 91 wherein said polypeptide comprises an amino acid sequence selected from the group consisting of: SEQ ID NO: 74; SEQ ID NO: 186; SEQ ID NO:78; SEQ ID NO:80; SEQ ID NO:82; SEQ ID NO:84; SEQ ID NO: 86; SEQ ID NO:90; and SEQ ID NO:94.

20

96. An isolated nucleic acid molecule comprising a nucleotide sequence that encodes a polypeptide comprising an amino acid sequence homologous to SEQ ID NO:76, and fragments thereof; said nucleic acid molecule encoding at least a portion of nGPCR-5.

97. An isolated polypeptide encoded by a nucleic acid molecule of claim 96.

SEQUENCE LISTING

```
<110> Pharmacia & Upjohn Company
        Vogeli, Gabriel
        Huff, Rita
        Sejlitz, Torsten
        Lind, Peter
        Slightom, Jerry
Schellin, Kathleen
        Bannigan, Chris
        Ruff, Valerie
        Kaytes, Paul
        Wood, Linda
        Parodi, Luis
        Hiebsch, Ronald
 <120> Novel G Protein Coupled Receptors
 <130> 043P1FHRM296
 <150> 60/165,838
 <151> 1999-11-16
 <150> 60/198,568
 <151 > 2000-04-20
<150> 60/166,071
:151 > 1999-11-17
4150 - 60/166,678
<151 / 1999-11-19
·150 · 60/173,396
<151> 1999-12-18
<150 - 60/184,129
<151 - 2000-02-22</pre>
<150> 60/185,401
<151 2000-02-28</pre>
+150 > 60/185,554
151 2000-02-28
<150. 60/186,530
151: 2000-03-02
+150> 60/186,811
<151> 2000-03-03
150> 60/188,114
151> 2000-03-09
· 150>
       60/190,310
151: 2000-03-17
· 150> 60/190,800
```

Page 1

151> 2000-03-21

```
<150×
       60/201,190
:151> 2000-05-02
<150>
       60/203,111
       2000-05-08
<151>
<150>
       60/207.094
<151> 2000-05-25
.160 > 190
170> PatentIn version 3.0
<210 > 1
<211>
      1182
-212> DNA
<213> H.Sapiens
<400 → 1
atotgggggt gggggatgct gggacagggg toaattgcot gaagcaagtg ctctcatccc
                                                                       50
cctagetect getgatetag ttggggetec agagtgggga ggagaaagge aetttgaaae
                                                                      120
ttetetgeee ttacegtett agecatcaaa etetgagetg gagatagtga egatgtgaca
                                                                      180
agaactttcc ctgggcctct ctgggccaca attcctggcc gagagaaaga ggaggaatga
                                                                      240
agtgagcace ticticacie etagggecat giggtagage igeaglegea estecticig
                                                                      3:10
ccaataggca tagatgagtg ggttgagcag ggagttgccc acqccgaqca gccacaggta
                                                                      350
cogttocago actaggtaga ggtgacacto otggcaggco acctgcacaa tgccagtgat
                                                                      4 20
aaqqaaqqqq qtccaqqata qaqcaaaqct cccaatqaqa acagacacag tacggagagc
                                                                      480
tttgaagteg etgggagtee gtggggateg ataaceteea gecatggete etgeatgtte
                                                                      540
catetttega atetgetgge tgtgcatgga ggcaatettg ageatgtege agtagaagaa
                                                                      600
qacaaagagg agcatggctg ggaagaagcc aacgcaggag agggtcagca cgaagtgagg
                                                                      660
otqaaataca gcaaagaago tgcactgccc tttgtaggca gtctgstgga acatggggat
                                                                      720
tocqaqtqgg aggaaqccaa tgaggtaaga cactaaccac agcccggcaa tgcaggccc
                                                                      780
ggccacgaac ccactcatga tottcaagta goggaagggo tgottgatgg caaggtacot
                                                                      840
etcaaaggtg atcagcatga cogtgaggac agaggcaget geggaggaag tgacaaatge
                                                                      900
cateegeagg stgcacaggg tettetgtgt gggccgagaa gggctggaga getggtetgt
                                                                      960
gagtaggeca gagatggeca caccaatcaa ggtgtcagec acagecagat tcaaggtgaa
                                                                     1020
gcagagactg acaccateat tettgtggat caacagcage acagecacag ccaetagtgt
                                                                     1080
gttagtagca atgatgaggg aggccaggac agcaaggatc actccaaatg agaaagatga
                                                                     1140
ttccatgtct cgaagtggca ggacttcact taccagggca tq
                                                                     1182
```

Page 2

<210> 2 <211> 335 <212> PRT <213> H.Sapiens <400> 2 Met Glu Ser Ser Phe Ser Phe Gly Val Ile Leu Ala Val Leu Ala Ser Leu Ile Ile Ala Thr Asn Thr Leu Val Ala Val Ala Val Leu Leu Leu Ile His Lys Asn Asp Gly Val Ser Leu Cys Phe Thr Leu Asn Leu Ala Val Ala Asp Thr Leu Ile Gly Val Ala Ile Ser Gly Leu Leu Thr Asp Gln Leu Ser Ser Pro Ser Arg Pro Thr Gln Lys Thr Leu Cys Ser Leu Arg Met Ala Phe Val Thr Ser Ser Ala Ala Ala Ser Val Leu Thr Val Met Leu Ile Thr Phe Asp Arg Tyr Leu Ala Ile Lys Gln Pro Phe Arg Tyr Leu Lys Ile Met Ser Gly Phe Val Ala Gly Ala Cys Ile Ala Gly Leu Trp Leu Val Ser Tyr Leu Ile Gly Phe Leu Pro Leu Gly Ile Pro Met Phe Gln Gln Thr Ala Tyr Lys Gly Gln Cys Ser Phe Phe Ala Val 150 Phe His Pro His Phe Val Leu Thr Leu Ser Cys Val Gly Phe Phe Pro Ala Met Leu Leu Phe Val Phe Phe Tyr Cys Asp Met Leu Lys Ile Ala Ser Met His Ser Gln Gln Ile Arg Lys Met Glu His Ala Gly Ala Met Ala Gly Gly Tyr Arg Ser Pro Arg Thr Pro Ser Asp Phe Lys Ala Leu Arg Thr Val Ser Val Leu Ile Gly Ser Phe Ala Leu Ser Trp Thr Pro Phe Leu Ile Thr Gly Ile Val Gln Val Ala Cys Gln Glu Cys His Leu Tyr Leu Val Leu Glu Arg Tyr Leu Trp Leu Leu Gly Val Gly Asn Ser

Leu Leu Asn Pro Leu Ile Tyr Ala Tyr Trp Gln Lys Glu Val Arg Leu 275 280 285 Gln Leu Tyr His Met Ala Leu Gly Val Lys Lys Val Leu Thr Ser Phe Leu Leu Phe Leu Ser Ala Arg Asn Cys Gly Pro Glu Arg Pro Arg Glu Ser Ser Cys His Ile Val Thr Ile Ser Ser Ser Glu Phe Asp Gly <210> 3 <211> 657

<212> DNA <213> H.Sapiens

<400> 3 caqcqcqaqc qccttcatgg tgacggtgtc catgcgctgg cagtgtctgc gtgccacccq 60 atqcacctqq aqcqaqqtqa qqcaqaqcac cqccaqcqqc aqcacqaaqc ccacqqcatq 120 gageqtqqeq qtqaaqqetq eqaaqeqeqq acqctcaqqe teqqqeqqea qqeqcaqeqa 180 acaqqacqcq aaqqcqctqc tgtaqccaaq ccacqaqcaq ccaaqtqcaq cqcctqaqaa 240 grocagegae tgtccccagg cacageccag cagcaggeeg geatagegeg gtegeaggeg 300 teeqqeqtaq eqeaqtqqqa ageecaetge cagecaetgg tetqegetca gegeeqeeae 360 getcagegec gegttggaeg ccaggaaggt gtccaggaag ccaatgaett ggcatgegee 420 gageqeeqae qqtqteeqee egeqeateae acegageage qtqaaqqqea tqteeaqeqe 480 540 enceageage aggtggeeca gagacagatt caccaggagg acgoetgagg etegagtgeg gageteageg etgtaggege aacaaageag caccagtgeg ttggatagea gegecaegge 600 magtaccate accaggagae cogcoagoag ogcotogoog gggocoatgg ogctago 657

<:10> 4 <211> 217 <212> PRT <213> H.Sapiens

<400>

Ser Ala Met Gly Pro Gly Glu Ala Leu Leu Ala Gly Leu Leu Val Met

Val Leu Ala Val Ala Leu Leu Ser Asn Ala Leu Val Leu Leu Cys Cys

Ala Tyr Ser Ala Glu Leu Arg Thr Arg Ala Ser Gly Val Leu Leu Val 40 45

Asn Leu Ser Leu Gly His Leu Leu Leu Ala Ala Leu Asp Met Pro Phe 55

Thr 65	Leu	Leu	Gly	Val	Met 70	Arg	Gly	Arg	Thr	Pro 75	Ser	Ala	Pro	Gly	Ala 80	
Суз	Gln	Val	Ile	Gly 85	Phe	Leu	Asp	Thr	Phe 90	Leu	Ala	Ser	Asn	Ala 95	Ala	
Leu	Ser	Val	Ala 100	Ala	Leu	Ser	Ala	Asp 105	Gln	Trp	Leu	Ala	Val 110	Gly	Phe	
F·ro	Leu	Arg 115	Tyr	Ala	Gly	Arg	Leu 120	Arg	Pro	Arg	Tyr	Ala 125	Gly	Leu	Leu	
Leu	Gly 130	Cys	Ala	Trp	Gly	Gln 135	Ser	Leu	Ala	Phe	Ser 140	Gly	Ala	Ala	Leu	
Gly 145	Cys	Ser	Trp	Leu	Gly 150	Tyr	Ser	Ser	Ala	Phe 155	Ala	Ser	Cys	Ser	Leu 160	
Arg	Leu	Pro	Pro	Glu 165	Pro	Glu	Arg	Pro	Arg 170	Phe	Ala	Ala	Phe	Thr 175	Ala	
Thr	Leu	His	Ala 180	Val	Gly	Phe	Val	Leu 185	Pro	Leu	Ala	Val	Leu 190	Cys	Leu	
Thr	Ser	Leu 195	Gln	Val	His	Arg	Val 200	Ala	Arg	Arg	His	Cys 205	Gln	Arg	Met	
Asp	Thr 210	Val	Thr	Met	Lys	Ala 215	Leu	Ala								
-212 -212 -212	1> 2 2> 1	5 222 DNA H.Sap	piens	3												
· 400		5 ata 1	rgato	nt og	at to	cctt	tatao	c ato	cata	caca	cact	cat t	oga ,	atgg	gatttt	60
			_												tctgta	120
tati	aaca	ttg :	tcata	cate	ag c	tatg	atcg	a ta	cctg	tcag	tct	caaa	tgc	tgta	agtcga	180
aca	catt.	aat 1	ttato	cccc	ct ta	agaa	gatt	a tgi	taaa	tgta	ta					222
-21: -21: -21: -21:	1> 2>	6 73 PRT H.Sa	pien:	5												
-:40	0:>	6														
Cys l	Ala	Gly	Val	Ile 5	Ser	Ile	Pro	Leu	Tyr 10	Ile	Pro	His	Thr	Leu 15	Phe	
Glu	Trp	Asp	Phe 20	Gly	Lys	Glu	Ile	Cys 25	Val	Phe	Trp	Leu	Thr 30	Thr	Asp	

Tyr Leu Leu Cys Thr Ala Ser Val Tyr Asn Ile Val Leu Ile Ser Tyr

Asp Arg Tyr Leu Ser Val Ser Asn Ala Val Ser Arg Thr His Phe Ilc

Pro Leu Arg Arg Leu Cys Lys Cys Ile 65 70

<210> 7

<211> 507 <212> DNA <213> H.Sapiens

<4005 7

<400> 7						
gacgtcgaag	caggtgatga	tgcccagggc	gtgcaccggg	taggtgagat	cggtgcgcgc	60
cagcggggac	agggcggtca	ggagcagcag	ccaggtccct	gcacacgcgg	ccaccgcgta	120
acgacggcgg	cgccagcgct	tggagctgag	cgggtacagg	atccccagga	agcgctccac	180
gctgatacag	gtcatggtga	ggatgctgga	atacatgttt	gcgtaaaagg	ccacggtcac	240
cacgttgcaa	agcagcaccc	cgaataccca	gtggtggcgg	ttgcaatggt	agtagatttg	300
gaaaggcaac	acgctggcca	gcatcaggtc	cgtgacgctc	aggttgatca	tgaagatgac	360
agacygggat	ctgggcccca	tgcgccggca	cagcacccac	agagagaaga	ggttgcccgg	420
gatgetgace	gccgccacca	gcgagtacac	cacgggcagg	gccaccgcga	tcgccgggtt	480
cogcagcato	tgcagcgtcg	cgttgtc				507

^{210&}gt; 8

-400: 8

Asp Asn Ala Thr Leu Gln Met Leu Arg Asn Pro Ala Ile Ala Val Ala

Leu Pro Val Val Tyr Ser Leu Val Ala Ala Val Ser Ile Pro Gly Asn

Leu Fhe Ser Leu Trp Val Leu Cys Arg Arg Met Gly Pro Arg Ser Pro

Ser Val Ile Phe Met Ile Asn Leu Ser Val Thr Asp Leu Met Leu Ala

Ser Val Leu Pro Phe Gln Ile Tyr Tyr His Cys Asn Arg His His Trp

Val Fhe Gly Val Leu Cys Asn Leu Val Val Thr Val Ala Phe Tyr Ala

^{/211&}gt; 169 -212> PRT -213> H.Sapiens

Asn	Met	Tyr	Ser 100	Ser	Ile	Leu	Thr	Met 105		Cys	Ile	Ser	Val 110	Glu	Arg	
Phe	Leu	Gly 115	Ile	Leu	Tyr	Pro	Leu 120		Ser	Lys	Arg	Trp 125	Arg	Arg	Arg	
Arg	Tyr 130	Ala	Val	Ala	Ala	Cys 135	Ala	Gly	Thr	Trp	Leu 140	Leu	Leu	Leu	Thr	
Ala 145		Ser	Pro	Leu	Ala 150	Arg	Thr	Asp	Leu	Thr 155		Pro	Val	His	Ala 160	
Leu	Gly	Ile	Ile	Thr 165	Cys	Phe	Asp	Val								
<21 <21 <31 <21	1 > 2 >	9 270 DNA H.Sa	pien:	s												
<40 ccc		9 tcc	tgcto	ectg	gg ca	agcet	cac	g tto	gtcg	gatc	tgci	tggca	agg (cqcc	geetae	60
															ggttc	120
gca	cggg.	agg (gaggo	egtet	t c	gtggd	cacto	c act	tgcgt	ccg	tgct	tgago	cat d	ectg	ggcate	180
gcg	ctgg	agc (gcago	cctca	ac ca	atggo	egege	e ago	3 9999	eccg	cgco	cagta	ctc o	cagto	ggggg	240
các	acgc	tgg (cgato	ggcaç	ge eq	gegge	cctg	3								270
<210 <210 <210 <210	1> : 2> :	10 90 PRT H.Sap	oiens	5							,					
<:400)> :	10														
Pro	Met	Phe	Leu	Leu 5	Leu	Gly	Ser	Leu	Thr 10	Leu	Ser	Asp	Leu	Leu 15	Ala	
Gly	Ala	Ala	Tyr 20	Ala	Ala	Asn	Ile	Leu 25	Leu	Ser	Gly	Pro	Leu 30	Thr	Leu	
Lys	Leu	Ser 35	Pro	Ala	Leu	Trp	Phe 40	Ala	Arg	Glu	Gly	Gly 45	Val	Phe	Val	
Ala	Leu 50	Thr	Ala	Ser	Val	Leu 55	Ser	Leu	Leu	Gly	Ile 60	Ala	Leu	Glu	Arg	
Ser 65	Leu	Thr	Met	Ala	Arg 70	Arg	Gly	Pro	Ala	Pro 75	Val	Ser	Ser	Arg	Gly 80	
	Thr	Leu	Ala	Met 85	Ala	Ala	Ala	Ala	Trp 90							
)>]	Leu 11 388	Ala		Ala	Ala	Ala	Ala								

PCT/US00/31581 WO 01/36473

<212> DNA <213> H.Sapiens

<400> 11 ctgctcattg	tggcctttgt	getgggegea	ctaggcaatg	gggtcgccct	gtgtggtttc	60
tgcttccaca	tgaagacctg	gaagcccagc	actgtttacc	ttttcaattt	ggccgtggct	120
gatttectee	ttatgatctg	cctgcctttt	cggacagact	attacctcag	acgtagacac	180
tgggcttttg	gggacattcc	ctgccgagtg	gggctcttca	cgttggccat	gaacagggcc	240
gggagcatcg	tgttccttac	ggtggtggct	gcggacaggt	atttcaaagt	ggtccacccc	300
caccacgcgg	tgaacactat	ctccacccgg	gtggcggctg	gcatcgtctg	caccctgtgg	360
gccctggtca	tectgggaac	agtgtatctt	ttgctggaga	accatctctg	cgtgcaagag	420
acggccgtct	cctgtgagag	cttcatcatg	gagteggeea	atggctggca	tgacatcatg	480
ttccagctgg	agttctttat	gcccctcggc	atcatcttat	tttgctcctt	caagattgtt	540
tqgagcctga	ggcggaggca	gcagctggcc	agacaggete	ggatgaagaa	ggcgacccgg	600
itcatcatgg	tggtggcaat	tgtgttcatc	acatgctacc	tgcccagcgt	gtctgctaga	660
ofctatttcc	totggacggt	gccctcgagt	gcctgcgatc	cctctgtcca	tggggccctg	720
dacataaccc	teagetteae	ctacatgaac	agcatgctgg	atcccctggt	gtattatttt	780
tuaagoccot	cctttcccaa	attctacaac	aagctcaaaa	tctgcagtct	gaaacccaag	840
cagocaggac	actcaaaaac	acaaaggccg	gaagagatgc	caatttcg		888

^{1110&}gt; 12

-:400> 12

Leu Leu Ile Val Ala Phe Val Leu Gly Ala Leu Gly Asn Gly Val Ala

Tyr Leu Phe Asn Leu Ala Val Ala Asp Phe Leu Leu Met Ile Cys Leu

Pro Phe Arg Thr Asp Tyr Tyr Leu Arg Arg Arg His Trp Ala Phe Gly 50 60

Asp Ile Pro Cys Arg Val Gly Leu Phe Thr Leu Ala Met Asn Arg Ala 65 70 75 80

Gly Ser Ile Val Phe Leu Thr Val Val Ala Ala Asp Arg Tyr Phe Lys 90

^{-211&}gt; 296 -2212> PRT

^{-213&}gt; H.Sapiens

Val	Val	His	Fro 100	His	His	Ala	Val	Asn 105	Thr	Ile	Ser	Thr	Arg 110	Val	Ala		
Ala	Gly	Ile 115	Val	Cys	Thr	Leu	Trp 120	Ala	Leu	Val	Ile	Leu 125	Gly	Thr	Val		
Tyr	Leu 130	Leu	Leu	Glu	Asn	His 135	Leu	Cys	Val	Gln	Glu 140	Thr	Ala	Val	Ser		
Cys 145	Glu	Ser	F'he	Ile	Met 150	Glu	Ser	Ala	Asn	Gly 155	Trp	His	Asp	Ile	Met 160		
Phe	Gln	Leu	Glu	Phe 165	Phe	Met	Pro	Leu	Gly 170	Ile	Ile	Leu	Phe	Cys 175	Ser		
Fhe	Lys	Ile	Val 180	Trp	Ser	Leu	Arg	Arg 185	Arg	Gln	Gln	Leu	Ala 190	Arg	Gln		
Ala	Arg	Met 195	Lys	Lys	Ala	Thr	Arg 200	Phe	Ile	Met	Val	Val 205	Ala	Ile	Val		
Phe	Ile 210	Thr	Cys	Tyr	Leu	Pro 215	Ser	Val	Ser	Ala	Arg 220	Leu	Tyr	Phe	Leu		
Trp 225	Thr	Val	Pro	Ser	Ser 230	Ala	Cys	Asp	Pro	Ser 235	Val	His	Gly	Ala	Leu 240		
His	Ile	Thr	Leu	Ser 245	Phe	Thr	Tyr	Met	Asn 250	Ser	Met	Leu	Asp	Pro 255	Leu		
Val	Tyr	Tyr	Phe 260	Ser	Ser	Pro	Ser	Phe 265	Pro	Lys	Pne	Tyr	Asn 270	Lys	Leu		
Lys	Ile	Cys 275	Ser	Leu	Lys	Pro	Lys 280	Gln	Pro	Gly	His	Ser 285	Lys	Thr	Gln		
Arg	Pro		Glu	Met	Pro	Ile 295	Ser										
·21 ·21 ·21	1 2	13 510 DNA H.Sa	pien	s													
<:40	0>	13															
															gtgcta		0
ti ti g	ccct	tac	tcat	catc	ac c	tact	cact	a ga	tgac	aggt	ggc	cctt	cgg	ggag	ctgctc	12	0
tgc	aagc	tgg	tgca	cttc	ct g	ttct	atat	c aa	cctt	tacg	gca	gcat	cct	gctg	ctgacc	18	0
tgc	atct	ctg	tgca	ccag	tt c	ctag	gtgt	g tg	ccac	ccac	tgt	gtto	gct	gccc	taccgg	24	0
acc	egga	ggc	atgo	ctgg	ct g	ggca	ccag	c ac	cacc	tggg	ccc	tggt	ggt	cctc	cagctg	30	0
ctq	jecca	cac	tggo	cttc	tc c	caca	cgga	c ta	cate	aatg	l ácc	agat	gat	ctgg	tatgac	36	0
atç	gacca	igad	aaga	gaat	tt t	gato	ggct	t tt	tgcc.	tacç Page		itagt	tct	gaca	ttgtct	42	0

WO 01/36473

PCT/US00/31581

caagccagag gagaacctca tgaggacagg												510				
<pre><210> 14 <211> 170 <212> PRT <213> H.Sapiens</pre>																
< 400)>	14														
Trp 1	Ser	Cys	Ala	Thr 5	Thr	Tyr	Leu	Val	Asn 10	Leu	Met	Val	Ala	Asp 15	Leu	
Leu	Tyr	Val	Leu 20	Leu	Pro	Phe	Leu	11e 25	Ile	Thr	Tyr	Ser	Leu 30	Asp	Asp	
Arg	Trp	Pro 35	Phe	Gly	Glu	Leu	Leu 40	Суѕ	Lys	Leu	Val	His 45	Phe	Leu	Phe	
17.r	Ile 50	Asn	Leu	Tyr	Gly	Ser 55	Ile	Leu	Leu	Leu	Thr 60	Cys	Ile	Ser	Val	
His 65	Gln	Phe	Leu	Gly	Val 70	Cys	His	Pro	Leu	Cys 75	Ser	Leu	Pro	Tyr	Arg 80	
Thr	Arg	Arg	His	Ala 85	Trp	Leu	Gly	Thr	Ser 90	Thr	Thr	Trp	Ala	Leu 95	Val	
Val	Leu	Gln	Leu 100	Leu	Pro	Thr	Leu	Ala 105	Phe	Ser	His	Thr	Asp 110	Tyr	Ile	
Asn	Gly	Gln 115	Met	Ile	Trp	Tyr	Asp 120	Met	Thr	Ser		Glu 125	Asn	Phe	Asp	
Arg	Leu 130	Phe	Ala	Tyr	Gly	Ile 135	Val	Leu	Thr	Leu	Ser 140	Gly	Phe	Leu	Ser	
Leu 145	Leu	Gly	His	Phe	Gly 150	Val	Leu	Phe	Thr	Asp 155	Gly	Gln	Glu	Pro	Asp 160	
Gln	Ala	Arg	Gly	Glu 165	Pro	His	Glu	Asp	Arg 170							
4210 4211 4212 4213	> 8	.5 194 DNA 1.Sap	iens													
r220 r 221 r 222 r 223	- π	nisc 431) i is	(4	61)	eoti	de										
400> 15 ccaectaegeg cageaegeeg acagggeete teeeteecat teteeegeag geeeggaega 60 Page 10											60					

ggetttettt eceteettgg teattitggt gtgetattea etgatggtea ggageetgat 480

120 ccacqctqcc tccaqccqqt cqqcaaacta qqqcaqctcq caqcccacqa acaqcaqccc caqcaqctqq ctcatcttca qqctctgcac cttggcgcqq qqcatcqcqc tqqqcqcacq 180 agetecacet aggstegeeg ascaggeegs tgsaceeget agggeettea geoggtgeeg 240 ccaccagacg gagagtaggt qqccacaaq: qasacccatg atottaacag gcqccacqaa 300 qeecqeqaeq qeetcataqa aegegtacae etgeaegtqe eageqetqea qqaqeqeqaa 360 matecaqtqq caqegaegea tecceggeea ggetegggeg gagagtggeg egeetggetg 420 cagagacqtt nnnnnnnnn nnnnnnnnn nnnnnnnnn nagtactage geaceacaa 480 540 necegações egegecagea geagtgecag cagecagede agggeggega gggcaegege aggcagegge eggeogtigeg gaagaegeac egilgegeegg egetegaggg egatgageac 600 cacqaqqtqq qooqaqqoqo occqeeegga tgeetgcaqo aqotqeagga ageggcaege 660 720 caqqtccccc qtqqccqcqc ggggctcgcc caqcaqttcc caqqccaqct qtqacaqcqc eqtqccccq cacqeqtaca qqtccqccaq qqccaqctqc accaqcaqqa aqtccatctt 780 qeqacqettn nnnnnnnnn nnnnnnnnn nnnnnnnae aggeggeaea geaetgtggt 840 attacctace accaseacea chaqqatqan ecchaqqaan accaqqoqqa eqeq 894

```
<2105 16
<2112 296
<2122 PRT
<2132 H.Sapiens

0202
<2212 UNSURE
0222 (26)..(35)
<2235 Xaa is unknown</pre>
```

<2208
-221> UNSURE
-222> (144)..(154)

-223> Xaa is Unknown

-400> 16

Arg Val Arg Leu Val Phe Leu Gly Val Ile Leu Val Val Ala Val Ala i 5 10 15

Xaa Xaa Lys Arg Arg Lys Met Asp Phe Leu Leu Val Gln Leu Ala 35 40 45

Leu Ala Asp Leu Tyr Ala Cys Gly Gly Thr Ala Leu Ser Gln Leu Ala Page 11

	50					55					60					
Trp 65	Glu	Leu	Leu	Gly	Glu 70	Pro	Arg	Ala	Ala	Thr 75	Gly	Asp	Leu	Ala	Cys 80	
Arg	Phe	Leu	Gln	Leu 85	Leu	Gln	Ala	Ser	Gly 90	Arg	Gly	Ala	Ser	Ala 95	His	
Leu	Val	Val	Leu 100	Ile	Ala	Leu	Glu	Arg 105	Arg	Arg	Ala	Val	Arg 110	Leu	Pro	
His	Gly	Arg 115	Pro	Leu	Pro	Ala	Arg 120	Ala	Leu	Ala	Ala	Leu 125	Gly	Trp	Leu	
Leu	Ala 130	Leu	Leu	Leu	Ala	Arg 135	Gly	Ser	Gly	Phe	Val 140	Val	Arg	Tyr	Xaa	
Хаа 145	Xaa	Xaa	Xaa	Xaa	Xaa 150	Xaa	Хаа	Xaa	Xaa	Thr 155	Ser	Leu	Gln	Pro	Gly 160	
Ala	Pro	Leu	Ser	Ala 165	Arg	Ala	Trp	Pro	Gly 170	Met	Arg	Arg	Cys	His 175	Trp	
Ile	Phe	Ala	Leu 180	Leu	Gln	Arg	Trp	His 185	Val	Gln	Val	Tyr	Ala 190	Phe	Tyr	
Glu	Ala	Val 195	Ala	Gly	Phe	Val	Ala 200	Pro	Val	Lys	Ile	Met 205	Gly	Val	Ala	
Cys	Gly 210	His	Leu	Leu	Ser	Val 215	Trp	Trp	Arg	His	Arg 220	Leu	Lys	Ala	Pro	
Ala 205	Gly	Ala	Ala	Ala	Trp 230	Ser	Ala	Ser	Pro	Gly 235	Gly	Ala	Arg	Ala	Pro 240	
Ser	Ala	Met	Pro	Arg 245	Ala	Lys	Val	Gln	Ser 250	Leu	Lys.	Met	Ser	Gln 255	Leu	
Leu	Gly	Leu	Leu 260	Phe	Val	Gly	Cys	Glu 265	Leu	Pro	Phe	Ala	Asp 270	Arg	Leu	
Glu	Ala	Ala 275	Trp	Ser	Ser	Gly	Pro 280	Ala	Gly	Glu	Trp	Glu 285	Gly	Glu	Ala	
Lou	Ser 290	Ala	Cys	Cys	Ala	Trp 295	Trp									
<210 <211 <212 <213	.> 8 !> €	.7 801 NA I.Sap	oiens	i												
<400 tcta		.7 .tt t	ctct	gaac	t tt	gago	ctgt	gaa	ıaaaa	igaa	ggga	itget	ge d	ctcaç	gccac	60
ссса	igcet	ag a	tact	cact	c to	gagto	rccat	gaç	ıgtaç	tag	agga	cact	ga t	gaca	gtcat	120
gggg	agga	igg t	agaa	tagg	ja aç	gagg	gtgad	ctq	gato	gatg	aaat	tgta	iga t	ccac	atggg	180

cttgatgacc	gtacaggtgg	ccgaacctgg	gaccagggac	ccattgggga	agtagtggaa	240
cttgatgcca	tggatgctgg	tgttgggcag	ggagaagagc	acggagaagc	cccagacgat	300
gccgaggatc	ctgagggccc	ggcgccgggt	gctctgcagt	ttggcgcgga	acgggtgtag	360
gatggccacg	tagcgctcca	cgctgacggt	ggtgatgctg	aggatggagg	cgaagcacac	420
ggtctcaaag	agggccgtct	tgaagtagca	gcccacgggc	ccgaacaaga	aagggtagtt	480
gcgccacatc	tcatagacct	ccaggggcat	tccaaggagc	aggaccagga	ggtcagagac	540
cgccaggctg	aagaggtagt	agttggtggg	cgtcttcata	gcctggtgct	gcagaatcac	600
caggcacacc	aggacattgc	caatgacccc	caccacaaaa	attggcacat	acaccacaga	660
cacggggagg	aagaagtggc	tgcgccgagg	tccgcagagg	aaggccagat	actcctcggt	720
gctgttcagg	tgtttctgga	atggatcttc	tagtttctgc	tggtagatcc	aggaagcatt	780
ctgaagtttt	tccatccctg	a				801
						801

<210> 18 <211> 249 <212> PRT <213> H.Sapiens

·400> 18

Ser Gly Met Glu Lys Leu Gln Asn Ala Ser Trp Ile Tyr Gln Gln Lys 1 10 15

Leu Glu Asp Pro Phe Gln Lys His Leu Asn Ser Thr Glu Glu Tyr Leu 20 25 30

Ala Phe Leu Cys Gly Pro Arg Arg Ser His Phe Phe Leu Pro Val Ser 35 40 45

Val Val Tyr Val Pro Ile Phe Val Val Gly Val Ile Gly Asn Val Leu

Val Cys Leu Val Ile Leu Gln His Gln Ala Met Lys Thr Pro Asn Thr 65 70 75 80

Tyr Tyr Leu Phe Ser Leu Ala Val Ser Asp Leu Leu Val Leu Leu Leu 95

Phe Gly Pro Val Gly Cys Tyr Phe Lys Thr Ala Leu Phe Glu Thr Val 115 120 125

Cys Phe Ala Ser Ile Leu Ser Ile Thr Thr Val Ser Val Glu Arg Tyr 130 135 140

Val Ala Ile Leu His Pro Phe Arg Ala Lys Leu Gln Ser Thr Arg Arg 145 \$150\$

60

Arg Ala Leu Arg Ile Leu Gly Ile Val Trp Gly Phe Ser Val Leu Phe 165 Ser Leu Pro Asn Thr Ser Ile His Gly Ile Lys Phe His Tyr Phe Pro Asn Gly Ser Leu Vai Pro Gly Ser Ala Thr Cys Thr Val Ile Lys Pro Met Trp Ile Tyr Asn Phe Ile Ile Glr. Val Thr Ser Phe Leu Phe Tyr Leu Leu Pro Met Thr Val Ile Ser Val Leu Tyr Tyr Leu Met Ala Leu 240 230 235 Arg Val Ser Ile Ala Gly Val Ala Gly 245 #10 · 13 :211 · 232 :0.12 · DNA 113 H.Sapiens $\pm 400 + 19$ Ir calagatga tittigetat egigeaaatt aliggatiit eeaaeteeat eigiaateee stig: ctatg cattlatgaa tgaaaactto aaaaaaaatg ttttgtctgc agtttgttat 120 180 tacátagtha ataaaacott ototocagoa caaaggoatg gaaattoagg aattacaatg 222 if gengaaga aagcaaagtt ttooctcaga gagaatccag tg ·:::10 · 20 3:11 = 73 0212 PRT 2213 H Sapiens (400 - 25) The Lys Met Ile Phe Ala Ile Val Glm Ile Ile Gly Phe Ser Asn Ser ile dys Asn Pro Ile Val Tyr Ala Phe Met Asn Glu Asn Phe Lys Lys Asn Val Leu Ser Ala Val Cys Tyr Cys Ile Val Asn Lys Thr Phe Ser Pro Ala 31n Arg His Gly Asn Ser Gly Ile Thr Met Met Arg Lys Lys 55 Ala Lys Phe Ser Leu Arg Glu Asn Pro <210 - 21

Page 14

<31. 447 €212 - DNA

PCT/US00/31581 WO 01/36473

60

120

180

240

300

360

420

447

•
<400: 21 queacageat geagetteet gragaattee actitigiett tigeactigaa gaagatgagg
tatotggtga ccaggatoac cacatagaat aggaaccgtg aggtacatgt ggatgtgcag
catgocacte acaaattige agaagggeag eccaaacate caagtetiet igaigaggia
gitcaagcga aatggcactg tcagcagaaa aacgctgtgg accaccacca agttaatgac
caccatggtg gtcactgacc gggtgttcat tttcaccagg aggaaaagaa tggaaatgac
ameeaccage eegecaataa geactatgaa gtagaggetg attaagtggg gtgteactat
a:gatege.ia gaggaattee tggaggtatt gtggeeagge atacttggga agteacetgg
aqgaqaaaaa gcaccagagt aactgac
+ 0100 22 +011 + 149 +012 + PRT +(213) + H.Sapiens
-1000 22
Val Ser Tyr Ser Gly Ala Phe Ser Pro Pro Gly Asp Phe Pro Ser Met 1 5 10 15
Pro Gly His Asn Thr Ser Arg Asn Ser Ser Cys Asp Pro Ile Val Thr 20 25 30
Fro His Leu Ile Ser Leu Tyr Phe Ile Val Leu Ile Gly Gly Leu Val 35 40 45
Gly Val Ile Ser Ile Leu Phe Leu Leu Val Lys Met Asn Thr Arg Ser 50 55 60
Val Thr Thr Met Ala Val Ile Asn Leu Val Val Val His Ser Val Phe 65 70 75 80
Leu Leu Thr Val Pro Phe Arg Leu Thr Tyr Leu Ile Lys Lys Thr Trp 85 90 95
Met Phe Gly Leu Pro Phe Cys Lys Phe Val Ser Ala Met Leu His Ile 100 105 110

lle Pro His Leu Gen Gen Val Gl
n Arg Gin Ser Gly Ile Leu Gl
n Lys 130 $$135\ \ \,$

Thr Ala Cys Cys Gly 145

<213: H.Sapiens

+210> 23 +211> 222

Page 15

8N5, x11 - W7: 1146479A2 - -

KO128 IMA KO138 E.Sapiens	
::400.	60
grtggcgcct gcaatogcac ctaccacaaa cttgaccggg ggcagggggg caggcccgcc	120
agogmacacg gtcagcagca ccagtocatt gcagagcacg gagagcaaca cgatggccca	180
pacqueeagg eggatgeece agettteaaa gaggtaetea ca	222
memocangg aggacgases agastasana pagga ana	
HD 10.4	
400 24	
Cys Glu Tyr Leu Phe Glu Ser Trp Gly fle Arg Leu Ala Val Trp Ala 10 15	
ile Val Leu Leu Ser Val Leu Cys Asn Gly Leu Val Leu Leu Thr Val 20 25 30	
Phe Ala Gly Gly Pro Ala Pro Leu Pro Pro Val Lys Phe Val Val Gly	
Ala Ile Ala Gly Ala Asn Thr Leu Thr Gly Ile Ser Cys Gly Leu Leu 50 55 60	
Ala Ser Mal Asp Ala Leu Thr Leu Mal Ser	
+:710 + 0.5 +:711 + 0.46 +:712 + 100A +:213 + 9.Sapiens	
≈ 100 · .5	50
unebloatea tetacaeget caccaacege gacetgegee acgegeteet gegeetggte	60
rightingguad godactootig oggnagagad odgagtiggot oddagdagtid ggogagogog	120
gritgigganit cogggggsot gogoogotgo otgooocogg godttgatgg gagottoago	180
pyctugguąc geteatogos coagogogas gggetggasa ceagoggoto cacaggoago	240
arcyst	246
.210 - 76 .311 - 22 -312 - PRT -213 - H.Sapiens	

Page 16

.:400 - :6

Asn Pro Ile Ile Tyr Thr Leu Thr Asn Arg Asp Leu Arg His Ala Leu Leu Arq Leu Val Cys Cys Gly Arg His Ser Cys Gly Arg Asp Pro Ser Gly Ser Gln Gln Ser Ala Ser Ala Ala Glu Ala Ser Gly Gly Leu Arg Arg Cys Leu Pro Pro Gly Leu Asp Gly Ser Phe Ser Gly Ser Glu Arg Ser Ser Fro Gln Arg Asp Gly Leu Asp Thr Ser Gly Ser Thr Gly Ser Pro Gly -.210.- 27 $\leq 2.14 \times -4.70$ V212% DNA V213% H.Sapiens - 2201. 221) misc_feature 221 (81)..(106) 223 - n is any nucleic acid -400 - 27 egtqaaqaac agcqccacca tqaccaqcat qtqcaccacq cqcqctctqc qccqcqauqc 60 regegggtee geagestest nnnnnnnnn nnnnnnnnn nnnnnntgge agagettgeg 120 egogatgong gogtacatga ceacgatgag egocagegge gecaggtaga tigtgegagaa 180 240 ragnacauto etgiagacco tececateco ettotogge cagecotoco agcaegaqta 300 zaqaqqqtaq qaqqqttqc qqqcqtccac catqaaqtqq tqctcctcac gqqtqacqqt raqoqtqarq qooqaqqqac acatgatgaq cagoqocagq qoocaqatga oqqoqatqqt 360 dangagenee ttooqoaqqq toaqottoto goggaaaqqq tgcacqatgo anoqqaacot 420 +210+ 28 +211+ 139 +212+ PFT 213. H.Sapiens

220-

221 UNSURE $\{104\}..(113)$

2230 Xaa is Unknown

1400> 25

Phe Arg Cys Ile Val His Pro Phe Arg Glu Lys Leu Thr Leu Arg Lys Page 17

1				5					10					15		
Ala	Leu	Val	Thr 20	Ile	Ala	Val	Ile	Trp 25	Ala	Leu	Ala	Leu	Leu 30	Ile	Met	
Cys	Pro	Ser 35	Ala	Val	Thr	Leu	Thr 40	Val	Thr	Arg	Glu	Glu 45	His	His	Phe	
Het.	Val 50	Asp	Ala	Arg	Asn	Arg 55	Ser	Tyr	Pro	Leu	Tyr 60	Ser	Суз	Trp	Glu	
Ala 65	Trp	Fro	Glu	Lys	Gly 70	Met	Arg	Arg	Val	Tyr 75	Thr	Thr	Val	Leu	Phe 80	
S∈r	His	lle	Tyr	Leu 85	Ala	Pro	Leu	Ala	Leu 90	Ile	Val	Val	Met	Tyr 95	Ala	
Arg	Ile	F.la	Arg 100	Lys	Leu	Cys	Xaa	Xaa 105	Xaa	Xaa	Xaa	Хаа	Хаа 110	Xaa	Xaa	
Хаа	Glu	<i>R</i> .la	Ala	Asp	Pro	Arg	Ala 120	Ser	Arg	Arg	Arg	Ala 125	Arg	Val	Val	
His	Met 130	Leu	Val	Met	Val	Ala 135	Leu	Phé	Phe	Thr						
:21 :21 :21 :31	1: 2:- 1	29 318 DNA H.Sap	pien:	5												
-(40) 10a): :::::::::::::::::::::::::::::::::::	29 gaga :	tqaq	taat:	ca q	geac	ttct	t ga	ggtc	ettg	ttg	agca	gga	agca	gacaat	60
															gtgggg	120
															ccagag	180
															cttgac	240
																300
					ee g	eegg	ctgg	c tg	catg	udda	LLC	eget.	gga	Cacc	cagcag	
ggt	tagt	gạc	atgg	gece												318
<21 <21 <21 <21	1 · 2 ·	30 106 PRT H Sa	pien	s												
:40	0 .	30														
Gly 1	Pro	Met	Pro	Pro 5	Thr	Leu	Leu	Gly	11e 10	Arg	Gln	Asn	Gly	His 15	Ala	
4la	Ser	Arg	Arg 20	Leu	Leu	Gly	Met	Asp 25	Glu	Val	I.ys	Gly	Glu 30	Lys	Gln	
Leu	З1у	Arg	Met	Phe	Tyr	Ala	Ile	Thr	Leu	Leu	Phe	Leu	Leu	Leu	Trp	

35 40 45

Ser Pro Tyr Ile Val Ala Cys Tyr Trp Arg Val Phe Val Lys Ala Cys 50 60

Ala Val Pro His Arg Tyr Leu Ala Thr Ala Val Trp Met Ser Phe Ala 65 70 75 80

Gln Ala Ala Val Asn Pro Ile Val Cys Phe Leu Leu Asn Lys Asp Leu 85 90 95

Lys Lys Cys Leu Arg Thr His Ala Pro Cys 100 105

- -0210> 31
- +211> 354 +312> DNA
- 43135 H.Sapiens
- < 4000 31

tattetgtaa tgaagaatgt cattcacact geeattggca catccagtgg cetcacetag 60
mattqtgaaa geeetteggt tggtgtattg ceaetteatt ttaaaaggat geacaagtee 120
etggtgeett tecacageaa tgcaggteat agtgaggatt tetgteacaa cageggtaga 180
etggacaaat ggcaccatet tgcaaatgaa agcacetgca gtaaggaaat aggataaate 240
atacatcaaa acaaaaagaa taaaggttte atetgtgtet ttgtaattat cactatcagt 300
emattetgag cetetgecaa aaagtttgat aattgtaatt actetgtaga caca 354

- 4210. 32
- · .:117: 117
- -12121 PRT
- 32133 H Sapiens
- <400. 32

Val Tyr Arg Val Ile Thr Ile Ile Lys Lou Phe Gly Arg Gly Ser Glu l 5 10 15

Trp Thr Asp Ser Asp Asn Tyr Lys Asp Thr Asp Glu Thr Phe Ile Leu 20 25 30

Phe Val Leu Met Tyr Asp Leu Ser Tyr Phe Leu Thr Ala Gly Ala Phe

Ile Leu Thr Met Thr Cys Ile Ala Val Glu Arg His Gln Gly Leu Val

His Pro Phe Lys Met Lys Trp Gln Tyr Thr Asn Arg Arg Ala Phe Thr 85 90 95

Met Leu Gly Glu Ala Thr Gly Cys Ala Asn Gly Ser Val Asn Asp Ile Page 19

110

100 105

Leu His Tyr Arg Ile

<210 · 33 <211 · 621 <212 · DNA

<213 - H.Sapiens

<400 - 33 gagemacatg atetttttga agtäettgae ggtgtegtte ttgaeggtea egaageaeag 60 aqtqttgatc atgctgttgc tcatggcgat gcactcgacg atgtagaagg cagtgaggta 120 gigoriotec ticacaaaca oggiggggaa gaagiogogo acgaiggiga agoogiagaa 180 gagoaccag catagoacgt aggoggtgag gatgoacatg agcaccagga segtetteet 240 grigghagege ageotottge ggatotgeto tgtotggaat coagggaceg bottgaacca 300 quigo ocogg gagatoctgg catagoacag ggtoatggtg accangggge ocacgaatto 360 420 tutqicaaag ataaagagga agtaggastt gtagtagage tgetigteea baggeeagat etgquegeag aagatetttt eetggetett gacaatgaeg aggabegtet eggtggtgaa 480 gragueggaa gggatggega teaggatgga cacegteeas accallggeaa teaggeeagt 500 qqotqtttqq cactteatte gtggteteag eggatggaca atagecagat seetagggea 621 a madadaag tggaggdagd c

<...10 - 34

<d11 + 207
<d12 + PRT</pre>

<213 · H.Sapiens

<400 + 34

Gry Cys Leu His Leu Cys Ser Cys Pro Arg Tyr Leu Ala Ile Val His 1 10 15

Pro Leu Arg Pro Arg Met Lys Cys Gln Thr Ala Thr 3ly Leu Ile Ala

Low Yal Trp Thr Val Ser Ile Lew Ile Ala Ile Pro Sor Ala Tyr Phe

For Phr Glu Thr Val Leu Val Ile Val Lys Ser Glo Glu Lys Ile Phe

Cys Gly Gln Jle Trp Pro Val Asp Gln Gln Leu Tyr Tyr Lys Ser Tyr 65 70 75 80

Phe Leu Phe Ile Phe Gly Ile Glu Phe Val Gly Pro Val Val Thr Met 85 - 90 - 95

WO 01/36473	PCT/US00/31581
-------------	----------------

Thr Leu Cys Tyr Ala Arg Ile Ser Arg Glu Leu Trp Phe Lys Ala Val	
Pro Gly Phe Gln Thr Glu Gln Ile Arg Lys Arg Leu Arg Cys Arg Arg 115 120 125	
Lys Thr Val Leu Val Leu Met Cys Ile Leu Thr Ala Tyr Val Leu Cys 130 135 140	
Trp Ala Pro Phe Tyr Gly Phe Thr Ile Val Arg Asp Phe Phe Pro Thr 145 150 150	
Val Phe Val Lys Glu Lys His Tyr Leu Thr Ala Phe Tyr Ile Val Glu 165 170 175	
Cys Ile Ala Met Ser Asn Ser Met Ile Asn Thr Leu Cys Phe Val Thr 180 185 190	
Val Lys Asn Asp Thr Val Lys Tyr Phe Lys Lys Ile Met Leu Leu 195 200 205	
+210 + 35 +211 + 483 +212 + DNA +213 + H.Sapiens	
-400 - 35 Ragonacact goagtgatga aatcaaatgt ccaacaccaa ccatagtcac cattactaac	·50
laaghaqoca caaaacttoo ottooagggt gttoagcago agggacaggg cocagggcag	120
decaracate acaetteaca entitottee ecapcaecae caetaccae taggoogcae	180
quesquesagg cagcacteag tactgatggs actsageatg steaggesta saaggtaggs	240
anaquiteate acqetqqtqa aqaaqetaqq qaaattqatq qaqatqqaac aqaaqaaqtt	300
actgaggtac accaggoaat ttataatotg gaagoagagg aagaggaagt oggoooggo	360
caggetgadg aegtagacag agaaggegtt eetgegeatg eggaageeea ggageeagag	420
nacasacong tittoctacca geocgaccag ggoaatgaaa aggatcagga agacogggat	480
dag	483
+210;- 36	

-210:- 36 -211:- 161 -212: PET -213:- H.Sapiens

-400> 36

Lou IIe Pro Val Phe Leu IIe Leu Phe IIe Ala Leu Val Gly Leu Val . $$10\$

Gly Asn Gly Phe Val Leu Trp Leu Leu Gly Phe Arg Mct Arg Asn 20 \$25\$

Ala Phe Ser Val Tyr Val Leu Ser Leu Ala Gly Ala Asp Phe Leu Phe 35 40 45	
Leu Cys Phe Gin Ile lle Asn Cys Leu Val Tyr Leu Ser Asn Phe Phe 50 60	
Cys Ger Ile Ser Ile Asn Phe Pro Ser Phe Phe Thr Ser Val Met Thr 70 75 80	
Phe Ala Tyr Leu Val Gly Leu Ser Met Leu Ser Ala Ile Ser Thr Glu 85 90 95	
Cys Cys Leu Ser Val Leu Arg Pro Ile Trp Tyr Cys Cys Cys Cys Pro 100 105 110	
Arg Asn Leu Ser Thr Val Met Cys Ala Leu Pro Trp Ala Leu Ser Leu 115 120 125	
Leu Leu Asn Thr Leu Glu Gly Lys Phe Cys Gly Phe Leu Val Ser Asn 130 135 140	
Giy Asp Tyr Gly Trp Cys Trp Thr Phe Asp Phe Ile Thr Ala Val Trp 145 150 155 160	
≟eu	
<pre><210: 37 <211: 330 <211: DNA <13: H.Sapiens</pre>	
.400). 37 gagagtetga tiotgaetta cateacatat giaggeoigg geatitetäi tigeageoig	60
Aftectttget tgteegttga ggteetagte tggageeaag tgacaaagae agagateaee 1	20
tattracgce atgtgtgcat tgttaacatt geagecactt tgctgatgge agatgtgtgg 1	180
ticaritgtyg offectifct tagtygodda afaacadad acaagggafg fgfggdagod 2	240
acatrititing gloatitiett tracettict glatititet ggalgetige eaaggeacte 3	300
oftatectet atggaateat gattgtttte 3	30
<pre> <pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre>	
<4001 38	
Glu Mer Leu Ile Leu Thr Tyr Ite Thr Tyr Val Gly Leu Gly Ile Ser 1 5 10 15	
lle Cys Ser Leu Ile Lou Cys Leu Ser Val Glu Val Leu Val Trp Ser 20 25 30	

Gln Val Thr Lys Thr Glu lle Thr Tyr Leu Arg His Val Cys Ile Val 35 40 45

Asn Ile Ala Ala Thr Leu Leu Met Ala Asp Val Trp Phe Ile Val Ala 50 60

Ger Phe Leu Ser Gly Pro Ile Thr His His Lys Gly Cys Val Ala Ala 35 70 75 80

Thr Phe Phe Gly His Phe Phe Tyr Leu Ser Val Phe Phe Trp Met Leu 85 90 95

Ala Lys Ala Leu Leu Ile Leu Tyr Gly Ile Met Ile Val Phe $100 \,$ $105 \,$ $110 \,$

<210> 39

-:2118 628

:.'12: DNA

::213: H.Sapiens

4400 - 39

ttgtgtggca gtagagagat gtcaggcttc agagtcaaca agaactggat ttcaaactgg 60 atttgaggae ecceaeettt ggtaagtgae ttattatetg egageetetg tttetetet 120 ctttaaatga ggacagtaaa toocatacgg cagggtggtg gggagaatca gagatgatas 1.80 auctggtgat cacatotggt ttgtgttoop aggggoacca gactagggtt totgagoatg 24 1 gatomaaccg toccagtott oggtacaaaa otgacaccaa toaacggacg tgaggagact 300 cuttoctaca atcagaccot gagottoacy gtgotgacgt geatcattto cottgtogga 3611 itgacaggaa acgoggtagt gototggoto otgggotaco gbatgogoag gaabgotgto 4220 tucatotaça toctoaacot ggoogcagoa gaottoctot tectoagott coagattata 480 eqttequeat tacqueteat caatateags dateteatoo quaaaateet eqtitetqtq 5.00 atgacettic ectactitae aggeotgagi atgetgageg ecateageae egaqegetge 600 stgtctgttc tgtggcccat ctggtacc 628

· 0100 40

-1112 205

· 012. PET

· 113: H.Sapiens

- 400: 40

Leu Cys Gly Ser Arg Glu Met Ser Gly Phe Arg Val Asn Lys Asn Trp $\frac{1}{2}$ $\frac{1}{2}$ $\frac{1}{2}$

:le Ser Asn Trp Ile GJy Pro Pro Pro Leu Val Ser Asp Leu Leu Ser
20 25 30

Ala Ser Leu Cys Phe Ser Leu Leu Met Arg Thr Val Asn Pro Ile Arg 35 40 45

Glm Gly Gly G	Gly Glu Asn	Gln Arg 55	Tyr Ser	Trp Ser	His Le	ı Val Cys	
Val Pro Arg 0 65	Sly Thr Arg 70	Leu Gly	Phe Leu	Ser Met 75	: Asp Pro	Thr Val 80	
Pro Val Phe G	Gly Thr Lys 85	Leu Thr	Pro Ile 90	Asn Gly	/ Arg Glu	Glu Thr 95	
Pro Cys Tyr A	Asn Gln Thr 100	Leu Ser	Phe Thr 105	Val Leu	Thr Cys		
Ser Leu Val G	Sly Leu Thr	Gly Asn 120	Ala Val	Val Leu	Trp Let 125	: Leu Gly	
Tyr Arg Met A	Arg Arg Asn	Ala Val 135	Ser Ile	Tyr Ile		n Leu Ala	
Ala Ala Asp E 145	Phe Leu Phe 150	Leu Ser	Phe Gln	Ile Ile 155	e Arg Sei	Pro Leu 160	
Aig Leu Ile A	Asn Ile Ser 165	His Leu	Ile Arg 170	Lys Ile	e Leu Vai	Ser Val 175	
Met Thr Phe E	Pro Tyr Phe 180	Thr Gly	Leu Ser 185	Met Let	ı Ser Ala 190		
Thr Glu Arg C 195	Cys Leu Ser	Val Leu 200	Trp Pro	Ile Trp	205		
-0010 - 41 -0110 - 319 -0112 - DNA -0113 - H.Sapi	iens						
<pre><400 41 arageaagca ac</pre>	mancancaa a	acettagge	: ataqtea	ataa aac	itatttat	attatactaa	60
ergovettet tt							120
etgtacaatg to	ctteatetg g	ctaggetat	: ttcaact	ctg ctt	tcaatcc:	cattttatat	180
goca:gcttt at	Locttagtt t	egcaaggea	ı ttgagga	atga tto	gtcacagg	catgatette	240
nacchtgäet dt	ttocacoot a	agcctgttt	totgod	catg ctt	aggctgt	gttcatcatt	300
daataggact ct	ttttctgg						319
0.10 42 0.11 103 0.12 PRT 0.13 H.Sapi	iens						

Page 24

Thr Slu Ser Lys Ala Thr Arg Thr Leu Gly Ile Val Met Gly Val Phe 1 5

Val Leu Cys Tro Leu Pro Phe Phe Val Leu Thr Ile Thr Asp Pro Phe Tie Asn Phe Thr Thr Leu Glu Asp Leu Tyr Asn Val Phe Leu Trp Leu Gly Tyr Phe Asn Ser Ala Phe Asn Pro Ile Leu Tyr Gly Met Leu Tyr Pro Trp Phe Arg Lys Ala Leu Arg Met Ile Val Thr Gly Met Ile Phe 55 His Pro Asp Ser Ser Thr Leu Ser Leu Phe Ser Ala His Ala Ala Val Phe Ile Ile Gln Asp Ser Phe 100 -:210:- 43 515 4:211:-:::12: DNA -3213 H.Sapiens 43 taqqaatoto aqaqaaqaaa gtaaggaaco agaaaaccat aaaagaatgt aaatggaaaa 60 quatcaqcaa atottattca ottatoacta aatotaaaat atgtoaaaat acatgaagac 120 aabayatqot ttäqaacaac tgttqaatqt attgtootac aacttggcat atgatcatgo 180 tracetetet atquecaaqt qtttattttt qeaqttqace ttaattteaa qttaqttttq 240 amptototae agtaatgttt ttaatotgto totaottott cagaaaataa attagttgtt 300 gaegaateag toottaagac ottgoogott acaataagtt ttattgoott occaaaccat 360 togtaaaaga aagcataaat caaggggtto atagotgaat tataataaac acaccaaact 420 admatctcat adacataagg aggagttata adattcatat aagcatcaat cactgcatca 480

-210: 44

1711: 148

-212: PRT

4213 H.Sapiens

argangtatg gtagccaaga gacaagaaat gotgo

<4000: 44

Pho Leu Val Ser Trp Leu Pro Tyr Leu Val Asp Ala Val Ile Asp Ala 20 25 30

Tyr Met Asn Phe Ile Thr Pro Pro Tyr Val Tyr Glu Ile Leu Val Trp 35 40 45

Page 25

515

Cys Mal Tyr Cyr Ash Se. 50	r Ala Met Asn 55	Pro Leu Ile Tyr Ala Phe Phe 60	
Tyr Gin Trp Phe Gly Ly 65 70	s Ala Ile Lys	Leu Ilo Val Ser Gly Lys Val 75 80	
Leu Arg Thr Asp Ser Se. 85	r Thr Thr Asn	Leu Phe Ser Glu Glu Val Glu 90 95	
Thr Asp Lys His Tyr Cy. 100	s Arg Asp Leu 105	Lys Thr Asn Leu Lys Leu Arg 110	
Ser Thr Ala Lys Ile As: 115	n Thr Trp Thr 120	Arg Gly Lys His Asp His Met 125	
Pro Ser Cys Arg Thr II	e His Ser Thr 135	Val Val Leu Lys His Leu Leu 140	
Ser Ser Cys Ile 145			
<pre><2100 45 <0.110 726 <0.120 DNA</pre>			
Clis. H.Sapiens			
cinos 45 ergguaagag gicologalo	tatectetae geo	egteetig gittigggge igigeiggea	60
grgtrtggaa acttactggt	catgatiget at	cottoact totaadaadt gcadadaddt	120
araaactttc tgattgcgtc	getggeetgt ge	tgacttot tggtgggagt cactgtgatg	180
d mithoagda qaqtqaqqta	tgtggagagc ty	ttggtact tiggggacag ttactgtaaa	240
ntocatabat gtittgacas	atotttotgt tt	tgettett tarrteattt atget@tate	300
nintgitgata gatacattgo	tgttactgat cc	totgacot atocaaccaa gtttactgtg	360
twagnttbag ggatatgcat	tgttetttee tg	gttotttt otgtoacata cagettttog	420
aretitiaea egggagecaa	cgaagaagga at:	tgaggaat tagtägttge tetaaeetgt	48)
graggagget gecaggetes	actgaatcaa aa	otgggted tactttgttt tettetatte	540
Litaladosa atgtogocat	ggtgtttata ta	cagtaaga taittttggt ggccaagcai	600
caggetagga agatagaaag	tacagecage ca	agotoagt cottotoaga gagttacaag	660
jiaaqaqtag caaaaagaga	gagaaagget ge	caaaacet tgggaattge tatggsagca	720
ritait			726

c:10 · 46
c:11 · 24;
<:12 · PRT
<213 > H.Sapiens

<400> 45

Leu Glu Arg Gly Pro Arg Ser Ile Leu Tyr Ala Val Leu Gly Phe Gly

Ala Val Leu Ala Ala Phe Gly Asn Leu Leu Val Met Ile Ala Ile Leu

His Phe Gln Leu His Thr Pro Thr Asn Phe Leu Ile Ala Ser Leu Ala

Cys Ala Asp Phe Leu Val Gly Val Thr Val Met Pro Phe Ser Thr Val

Arg Ser Val Glu Ser Cys Trp Tyr Phe Gly Asp Ser Tyr Cys Lys Phe

His Thr Cys Phe Asp Thr Ser Phe Cys Phe Ala Ser Leu Phe His Leu

Cys Cys Ile Ser Val Asp Arg Tyr Ile Ala Val Thr Asp Pro Leu Thr

Tyr Pro Thr Lys Phe Thr Val Ser Val Ser Gly Ile Cys Ile Val Leu

Ser Trp Phe Phe Ser Val Thr Tyr Ser Phe Ser Ile Phe Tyr Thr Gly 135

Ala Asn Glu Glu Gly Ile Glu Glu Leu Val Val Ala Leu Thr Cys Val

Gly Gly Cys Gln Ala Pro Leu Asn Gln Asn Trp Val Leu Leu Cys Phe 165

Leu Leu Phe Phe Ile Pro Asn Val Ala Met Val Phe Ile Tyr Ser Lys

Ile Phe Leu Val Ala Lys His Gln Ala Arg Lys Ile Glu Ser Thr Ala 200

Ser Gln Ala Gln Ser Phe Ser Glu Ser Tyr Lys Glu Arg Val Ala Lys

Arc Glu Arg Lys Ala Ala Lys Thr Leu Gly Ile Ala Met Ala Ala Phe

Leu

<2105 47 <2116 660

312 DNA

<213> H.Sapiens

<400: 47

alonging cottacted adjacency goottyteta typectitat daleagetyt 60

otcaatccag	ttctctatgt	cttcattggg	catgacttct	gggagcactt	getecaetee	120
ctgctagcig	cettagaacg	ggcacttagc	gaggagccag	atagtgcctg	aateccaget	180
cocaggdaga	tgagtccttt	ataacatgac	ccaatttcct	actocatttt	cccaccapts	240
aatoototto	ccaaacagct	ctaccataat	ccaacatcca	acagaattta	agagaataaa	300
ccacaactit	taagtgaget	ctatgtgcta	gglcatgttt	tagaatacaa	ccttaägtgc	360
ctqgaagatg	gaggcaagaa	acaaacaagg	totcattett	tagaggaaga	cagttcacca	420
agastsaaas	agaaaaaaag	atagttatct	tgtgacaaaa	caagtcataa	aattgggtca	480
ggacetgeag	caatgacttt	atgctagaat	ccagagcact	agcaggaaac	tgcttaaatt	540
ttacttaatc	aaagtcaagt	ttggacatac	atgtcaggta	aaacctagca	gagatgaget	600
accttgattt	taaaacttca	agggataget	caatgtcatc	aagatoottt	tgatgacttg	660

<210> 43

#2115 211 #2125 PRT #2135 H.Sapiens

<4000 43

Asn Gln Val Ala Leu Leu Leu Arg Pro Leu Ala Leu Ser Met Ala Phè

The Ash Ser Cys Leu Ash Pro Val Leu Tyr Val Phe The Gly His Asp

Phe Trp Giu His Leu Leu His Ser Leu Leu Ala Ala Leu Glu Arg Ala

Leu Ser Glu Glu Pro Asp Sor Ala Ile Pro Ala Pro Arg Gin Met Ser

Pro Led His Asp Pro Ile Ser Tyr Sor Ile Phe Pro Pro Leu Asn Pro

Leu Pro Lys Gln Leu Tyr His Asn Pro Thr Ser Asn Arg Ile Glu Asn

Lys Pro 31n Leu Leu Ser Glu Leu Tyr Val Leu Gly His Val Leu Glu

Tjr Asn Leu Lys Cys Leu Glu Asp Gly Gly Lys Lys Gln Thr Arg Ser

His Ser Leu Glu Glu Asp Ser Ser Pro Arg Leu Lys Gln Lys Lys Arg

Leu Ser Cys Asp Lys Thr Ser His Lys Ile Gly Ser Gly Pro Ala Ala

Met Thr Leu Cys Asn Pro Glu His Gln Glu Thr Ala Ile Leu Leu Asn Page 28

	165	170	175
Gln Ser Gln Val	Trp Thr Tyr Met Ser 185	Gly Lys Thr Gln Arg 190	Ala Thr
Leu Ile Leu Lys 195	Leu Gln Gly Ile Ala 200	Gln Cys His Gln Asp 205	Pro Phe
Asp Asp Leu			
<pre><:U10> 49 <u11> 465 <u12> DNA <u13 h.sapiens<="" pre=""></u13></u12></u11></pre>	5		
9400 - 49 gottatteac ggcca	accate cleaagetgt tge	ogcacijga ggaggogoac	ggcegggage 60
ageggaggeg egegg	gtggg: etggeegegg tge	statigat ggaattigta	acctgcttcg 120
o occaacaa ottog	gtgots otggogoada tog	gtgageeg eetgttetae	ggcaagaget 180
a staccacgt gtaca	agets aegetgtgts tea	igctgeet caacaactgt	etggaceegt 240
tigittatta citto	jegtes egggaattes age	etgegeet gegggaatat	ttgggetgee 300
geogggtgee cagaç	gacaco otggacacgo geo	gegagag ectettetee	gecaggacea 360
outergtgeg etec	gagges ggtgegeass etg	gaagggat ggagggagce	accaggonog 420
quetecagag geage	gagagt gtgttctgag tcc	cgggage geage	465
<pre>%: 10% 50 %:11% 160 %: 12% PRT %: 13% H.Sapiens</pre>	ŝ		
-:400h 50			
Lem Phe Thr Ala	Thr Ile Leu Lys Leu 5	Leu Arg Thr Glu Glu 10	Ala His 15
Gly Arg Clu Gln 20	Arg Arg Arg Ala Val 25	Gly Leu Ala Ala Val 30	Val Leu
Leu Ala Phe Val 35	Thr Cys Phe Ala Pro 40	Asn Asn Phe Val Leu 45	Leu Ala
His Ile Val Ser 50	Arg Leu Phe Tyr Gly 55	Lys Ser Tyr Tyr His 60	Val Tyr
Lys Leu Thr Leu 65	Cys Leu Ser Cys Leu 70	Asn Asn Cys Leu Asp 75	Pro Phe 80
Val Tyr Tyr Phe	Ala Ser Arg Glu Phe 85	Gln Leu Arg Leu Arg 90	Glu Tyr 95
		Page 20	

Page 29

BASE N. 1 - WE - 130473A2 9

603

, WO 01/36473

Leu Gly Cys Arg Arg Val Pro Arg Asp Thr Leu Asp Thr Arg Arg Glu 100 105 110	
Ser Leu Fne Ser Ala Arg Thr Thr Ser Val Arg Ser Glu Ala Gly Ala 115 120 125	
His Pro Glu Gly Met Glu Gly Ala Thr Arg Pro Gly Leu Gln Arg Gln 130 135 140	
Glu Ser Val Phe Val Pro Gly Ala Gln Ala Ala Pro Pro Gly Leu Arg 145 150 155 160	
12105 51 12115 603 12117 DNA 12137 H.Sapiens	
$(4.00 \leq 51)$. That the theoretical coalective attracetty coartest goodsattities.	60
stygg ofcat fittocctaft afsotycotc acattyatca agggatyayy otygcayyat	120
cognuaceca cagggeocog tgggcoùtga gaggeteetg gaettgaase teaggaeaet	180
iconcletgy ofgooggoay ggatggaago tggatgagda ggdaggagdt ggdagtgggg	240
jtgmagaged ataggetatt ggggtggada ggettgggtg deteatggga getessdatg	300
gradienting decetting detettattt eteacecean gettteeen gagaggttea	360
Aqtimgaaqa tgccccaaag atccacqtqq coctqqqtqq caqcctqttc ctcctqaatc	420
tagentente agecaatged aggagegest caaaaggggte egaegeegee egeegggess	480
gggggeret ettecactae tteetgetet gtgeetteae etggatggge ettgaagest	5 1 0
termineteta ectgeteget gteagg@tet teaacaceta ettegggcae tactt@etga	600

age

:01): 52 :01]: 198 :013: PET :213: H.Sapiens

:400> 52

No. The Tyr Ser Ala Leu Tyr Pro The Phe Ash Ser Leu Cys Tyr Ser I $$10^{\circ}$$

Prc Ala Ser Phe Ser Gly Leu Ile Phe Pro Ile Ile Leu Pro His Ile $20 \\ 25 \\ 30$

 $\mathsf{Asp}(\mathsf{Gln}|\mathsf{Gly}|\mathsf{Met}|\mathsf{Arg}|\mathsf{Leu}|\mathsf{Ala}|\mathsf{Gly}|\mathsf{Ser}|\mathsf{Gly}|\mathsf{Thr}|\mathsf{His}|\mathsf{Arg}|\mathsf{Ala}|\mathsf{Pro}|\mathsf{Trp}|_{\mathfrak{IS}}$

Ala Met Arg Gly Ser Trp Thr Thr Ser Gly His Ser His Ser Gly Cys 50 60Page 30

Arg Gln Gly Trp Lys Leu Asp Glu Gln Ala Gly Ala Gly Ser Gly Gly 65 70 80	
Gly Glu Pro Ala Ile Gly Val Asp Arg Leu Gly Cys Leu Met Gly Ala 85 90 95	
Pro His Gly Ser Cys Gly Pro Leu Gly Pro Leu Ile Ser His Pro Arg 100 105 110	
Leu Ser Arg Glu Arg Phe Lys Ser Glu Asp Ala Pro Lys Ile His Val 115 120 125	
Ala Leu Gly Gly Ser Leu Phe Leu Leu Asn Leu Ala Phe Leu Val Asn 130 135 140	
Val Gly Ser Gly Ser Lys Gly Ser Asp Ala Ala Cys Trp Ala Arg Gly 145 150 160	
Ala Val Phe His Tyr Phe Leu Leu Cys Ala Phe Thr Trp Met Gly Leu 165 170 175	
Glu Ala Phe His Leu Tyr Leu Leu Ala Val Arg Val Phe Asn Thr Tyr 180 185 190	
Phe Gly His Tyr Phe Leu 195	
+00100 53 +00110 335 +00120 DNA +00130 H.Sapiens	
:4000 53 Auttiggtegg agagtgeage tgettgaaat ggaggattga aateateace aggaggttte	60
omaanacago cagcacagoo ccamagocam acactatgta cagamtosco ogqqatocog	120
qugaqaaggg gattttcaca daggacccat toacgttogo gtagdacago tgdacagoda	180
nnagoaggga tgaattgotg otoataacgo tggtatttac atatggagaa attstigtoot	240
tottuattat cacaaaaaat acaggattgt tootgatttt cattgotoot googaaaaaa	300
acacatatto accaggatgo cagaggaaat gatca	335
1710% 54 1711% 111 1712% PRT 1713% H.Sapiens	

:400> 54

Asp His Phe Leu Trp His Pro Gly Glu Tyr Val Phe Phe Ser Ala Gly 1 $$ 15

Asp Lys Ile Ser Pro Tyr Val Asn Thr Ser Val Met Ser Ser Asn Ser 35 40 45	
Ser Leu Leu Va! Ala Val Gln Leu Cys Tyr Ala Asn Val Asn Gly Ser 50 55 60	
Cys Val Lys Ile Pro Phe Ser Pro Gly Ser Arg Val Ile Leu Tyr Ile 55 70 75 80	
Val Phe Gly Phe Gly Ala Val Leu Ala Val Phe Gly Asn Leu Leu Val 85 90 95	
Met Ile Ser Ile Lei His Phe Lys Gln Leu His Ser Pro Thr Asn . 100 105 110	
### ##################################	
<400: 55 cacarettaa caagaetgaa aaacattgat ttgtttttaa tttgaagage aatttatttg	-60
chatroatte atagtettae tigattitta aaaacteatt tegetiggia attitaaagg	1220
tateutgaan thogtotate caactgotta tatatgitoa gaaaacaaat toatgytigo	180
tqaantqtto tttaaaacct gabbagttab aataactttt attgotttoo taaaccatgg	240
jtaawataaa goataaatoa aaggattoat ggotgagtta taataagoac accaacagoa	300
toatmaatac aggoaggggt talaaagsoc ataaaggcat caattaatga atcaalgota	360
tutgutaaco atgadateat adatgetace aetgtgacee ecagggtttt agetgetttt	4.20
otototototo tggocactot ggotttgtaa otototgagg atgattotgt ottgotacca	430
gtatritteta tettiittege eigilegileta geeacaagaa ataigittaee alacagaati	540
icaraataa aygtaggtat aaagaaggat agaaaatotg toaaca	536
<pre><210 · 56 <2211 · 190 <212 · PRT <213 · H.Sapiens</pre>	
<100 · 56	
Lou Thr Asp Phe Leu Ser Phe Phe Ile Pro Thr Phe Ile Met Ile Ile 1 10 15	

Page 32

Glu Asn Thr Gly Ser Lys Thr Glu Ser Ser Ser Glu Ser Tyr Lys Ala $25 \,$

Arg Val Ala Arg Arg Glu Arg Lys Ala Ala Lys Thr Leu Gly Val Thr 50 60	
Val Val Ala Phe Met Ile Ser Trp Leu Pro Tyr Ser Ile Asp Ser Leu 55 70 75 80	
Ile Asp Ala Phe Met Gly Phe Ile Thr Pro Ala Cys Ile Tyr Glu Ile 85 90 95	
Cys Cys Trp Cys Ala Tyr Tyr Asn Ser Ala Met Asn Pro Leu Ile Tyr 100 105 110	
Ala Leu Phe Tyr Pro Trp Phe Arg Lys Ala Ile Lys Val Ile Val Thr 115 120 125	
Gay Gln Val Leu Lys Asn Ser Ser Ala Thr Met Asn Leu Phe Ser Glu 130 135 140	
His lie Ala Val Gly Thr Lys Phe Arg Ile Pro Leu Lys Leu Pro Ser 145 150 155 160	
Glu Met Ser Phe Lys Ser Ser Lys Thr Met Asn Glu Gln Ile Asn Cys 165 170 175	
Ser Ser Asn Lys Gln Ile Asn Val Phe Gln Ser Cys Asp Val 180 185 190	
12:10 57 12:11:- 976	
+2120 DNA +2130 H.Sapiens	
	60
213: H.Sapiens	60 120
213: H.Sapiens 400: 57 tttgtggdaa ggagaeddtg atdooggtdt tedtgatdet tttdattgdd etggtegggd	
213: H.Sapiens 400: 57 ttgtgggaa ggagaccetg atcocqgtet tectgateet tttcattgee etggteggge tgqtaggaaa egggtttgtg etetggetee tgggetteeg catgegeagg aacgeettet	120
213: H.Sapiens 400: 57 ttgrggdaa ggagadddg atddggdd tddgatdd tttdattgdd dtggdggd tggtaggaaa egggtttgrg ctdtggddd tgggdtteg batgdgaag aacgedtet dtgtdtaegt cdtaagodtg geoggggda aettactett cetetgette bagattataa	120 180
213: H.Sapiens 400: 57 *ttgtggdaa ggagaddotg atdddgdtd tddtgatdd tttdattgdd dtggtdggdd tggtuggaaa cgggtttgtg dtdtggdtdd tgggdttdg datgdgdagg aadgddtdt dtgtdtacgt ddtagddtg gddgggddg acttddtt ddtdgdtd dagattataa uttgddtggt gtaddtdagt aadttdttd gttdatdtd dataatttd datagdttdt	120 180 240
*213: H.Sapiens 400: 57 *ttgrgggaa ggagaccctg atcocqgtct toctgatcct tttcattgcc etggtcgggc tggtaggaaa cgggtttgtg ctctggctcc tgggcttccg catgcgcagg aacqccttct ctgtctacgt cctcagcctg geoggggccg acttcctctt cctctgcttc cagattataa attgcctggt gtacctcagt aacttcttct gttccatctc catcaatttc cctagcttct *cac.actgt gatgacctgt gcctaccttg caggcctgag catgctgagc accgtcagca	120 180 240 300
*213: H.Sapiens '400: 57 *ttgtgggaa ggagaccctg atcocqgtct toctgatcct tttcattgec ctggtcgggc tgqtuggaaa cgggtttgtg ctctggctcc tgggettccg catgcgcagg aacgcettct ctgtctacgt cctcagcctg geogggccg acttectett cctctgcttc cagattataa attgcctggt gtacctcagt aacttettet gttccatete catcaattte cctagettet reac,actgt gatgacctgt gcctaccttg caggcctgag catgctgage accgtcagea ccgaqcgctg cctgtccgtc ctgtggccca tctggtatcg ctgccgccgc cccagacacc	120 180 240 300 360
*213: H.Sapiens 400: 57 *ttgrgggaa ggagaccctg atcocqgtct toctgatcct tttcattgcc etggtegggc tggtaggaaa egggtttgtg etctggetce tgggetteeg batgegeagg aacqcettet etgtrtacgt octoagoetg geoggggeeg acttectett octotgette bagattataa attgeetggt gtabeleagt aacttettet gttebatete batgetgage accgeteet *cac.actgt gatgacctgt geotaccttg baggeetgag batgetgage accgeteagea *regagegetg betgteegte bagtggeeea tetggtateg baggeegege booggaeacc *gtbageggt batgtgtgtb bagetetggg beotgteeet actgetgage atcttggaag	120 180 240 300 360 420
*213: H.Sapiens *400* 57 *ttgtgggaa ggagacctg atcocqgtct tootgatcot titcattgcc otggtegggc tgqtaggaaa egggtttgtg ctotggetce tgggetteeg catgegeagg aacqcettet ctgtctacgt cotbagectg geogggeeg acttectett ectotgetce cagattataa attgcctggt gtacctcagt aacttettet gttcbatete catcaattte cotagettet reaclactgt gatgacctgt geotaccttg caggeetgag catgetgage accgtcagea regadegetg cotgtccgtc ctgtggeeca tetggtateg etgecgeege cocagacacc *gtcageggt egtgttgtc etgetetggg coctgtccct actgctgage atettggaag ggaacttetg tggettetta tttagtgatg gtgabtetgg ttggtgtcag abatttgatt	120 180 240 300 360 420 480
*213: H.Sapiens 400: 57 *ttgtgggaa ggagaccctg atcocqgtct toctgatcct tttcattgcc etggtegggc tggtaggaaa egggtttgtg ctctggctcc tgggettccg batgcgcagg aacqcettct etgtetacgt octoagoctg geoggggccg acttectett octctgettc bagattataa attgcetggt gtabeleagt aacttettet gttcoatetc batgattgage accgtcagea *cac.actgt gatgacctgt geotaccttg baggectgag batgctgage accgtcagea *cagaccgctg betgtecgtc bagtggccca tetggtateg baggecgcgc becagacacc *gtbagcggt baggtgtgtb baggecca batgctcct actgctgage acctgtgaag *ggaabttetg tggettctta tttagtgatg gtgabtctgg ttggtgtbag abatttgatt teatractgc agogtggotg attttttat beatjgttb baggggccc agtctggccc	120 180 240 300 360 420 480 540
*213: H.Sapiens *400: 57 *ttgtgggaa ggagacctg atcocqgtct toctgatcct titeattgcc etggtegggc tggtaggaaa egggtttgtg etetggetee tgggetteeg eatgegeagg aacqcettet etgtmtacgt ectsagectg geoggggeeg acttectett ectetgette eagattataa attgeetggt gtadeteagt aacttettet gttesatete eatoaattte estagettet reaclactgt gatgacctgt geotaccttg eaggetgag eatgetgage accgteagea megadegetg ectgteegte etgtggeeea tetggtateg etgeegeege escagacace *gteageggt egtgtgtte etgetetggg ecetgteeet actgetgage atettggaag ggaacttetg tggettetta titagtgatg gtgastetgg tiggtgteag asatttgatt teatmactge agegtggmtg attttttat teatigttet etgtgggtee agtetggeee tgetggteag gateetetgt ggeteeagg gtetjeeact gaccaggetg tacctgacca	120 180 240 300 360 420 480 540 600

840900960976

ttag	gaaç	gca q	gtggd	eggst	ig lea	agcac	eccga	ı tcc	ctcaa	gct	ggct	ctco	ag	agggo	ctctg
agga	catt	gc t	gagg	gtgga	at da	acagt	gaad	g gat	gatt	ccg	tcaç	gggc	cc	cggaq	gatte
aaga	agca	att d	etggt	gtag	gg ga	atgga	10000	tct	actt	.cca	tcat	atat	at	gtggd	itttg
gaggeaactt tgeece															
H210 - 58 H211 - 324 H212 PRT H213 - E.Sapiens H220 - H221 - UNSORE															
222		(266)	(2		ur.										
-1223		(aa :	is Ur	KHOV	VII										
0.400		3.5													
Cys	Gly	Lys	Glu	Thr 5	Leu	Ile	Pro	Val	Phe 10	Leu	Ile	Leu	Phe	11e 15	Ala
tioa	Val	G1 /	Leu 20	Val	Gly	Asn	Gly	Phe 25	Váil	Leu	Trp	Leu	Leu 30	Gly	Phe
Arg	let	Arg 35	Arg	Asn	Ala	Phe	Ser 40	Val	Туr	Val	Leu	Ser 45	Leu	Ala	Gly
	Asp 50	Phe	Len	Phe	Leı	Суs 55	Phe	Gln	Tle	Ϊle	Asn 60	Cys	Leu	Val	Туг
Lез 65	Ser	Asn	Phe	Phe	Суз 70	Ser	Ile	Ser	lle	Asn 75	Phe	Pro	Ser	Phe	Phe 80
Thr	Thr	Vei L	Met	Thr 85	Cys	Ala	Tyr	Leu	Ala 90	Gly	Leu	Ser	Met	Leu 95	Ser
Thr	Val	S∈r	Thr 100	Glu	Arj	Суѕ	Leu	Ser 105	Val	Leu	Trp	Pro	110	Trp	Tyr
Arg	Cys	Arg 115	Arg	Pro	Arg	His	Leu 120	Ser	Ala	Val	Val	Cys 125	Val	Leu	Leu
	Ala 130	Leu	Ser	Leu	Leu	Leu 135	Ser	Ile	Leu	Glu	Gly 140	Lys	Phe	Cys	Gly

Leg Thr Arg Leg Tyr Leg Thr Ile Leg Leg Thr Val Leg Val Ser Leg Page 34

Phe Leu Phe Ser Asp Gly Asp Ser Gly T:p dys Gln Thr Phe Asp Phe 145 $$ 155 $$ 160

Fig. Thr Ala Ala Trp Leu Ile Phe Leu Phe Met Val Leu Cys Gly Ser 165 . Ser Leu Ala Leu Leu Val Arg Ile Leu Cys Gly Ser Arg Gly Leu Pro

WO 01/36473	PCT/US00/31581

195	200	20	5
Leu Cys Gly Leu Pro Ph [10]	e Gly Ile Gln 215	Trp Phe Leu Il 220	e Leu Trp Ile
Trp Lys Asp Ser Asp Va 225 23		His Ile His Pr 235	o Val Ser Val 240
Val Leu Ser Ser Leu As 245	n Ser Ser Ala	Asn Pro Ile Il 250	e Tyr Phe Phe 255
Val Gly Ser Phe Arg Ly 260	s Gln Trp Arg 265	Xaa Gln His Pr	o Ile Leu Lys 270
Leu Ala Leu Gln Arg Al 275	a Leu Gln Asp 280	Ile Ala Glu Va 28	
Glu Gly Cys Phe Arg Gl 190	n Gly Thr Arg 295	Arg Phe Lys Gl 300	u Ala Phe Trp
Cys Arg Asp Gly Pro Le 305		His Ile Tyr Va 315	l Ala Leu Arg 320
Gly Asn Phe Ala			
+1110: 59 +1112: 578 +112: DNA +113: H.Sapiens			
·:400P 59			
entiquatet cactgitigag			
gtaanaggtt abcaaaggtg			
tggatotgat totoaatgga			
coctcaagat atggaagggt			
goottogtgo tttotgotta			
taatogtggt atagcaaagt	gtcactatea cea	aaggggag gcagaa	agta ettgeagtea 360
naatcaggtt gtancactta	atagtattga gtt	teatooga aetggt	gagg togagadagg 420
ctgatctgtt ggtoctgttg	gttgatgtga toa	aagaaqgt categg	aatg acagetacea 480
gigawatgat ocabaccaca	gcacaggeta esa	actgeaca tegagt	tttg tgaatggaaa — 540
ageageteat tgggtgaatg	atcacacagt ago	oggaag	578
<pre><210: 60 <211: 192 <212: PRT <213: H.Sapiens <400: 60</pre>			

Phe Arg Tyr Cys Val Ile Ile His Pro Met Ser Cys Phe Ser Ile H 1 5 10 15	is
Lys Thr Arg Cys Ala Val Val Ala Cys Ala Val Val Trp Ile Ile Se 20 25 30	er
Leu Val Ala Val Ile Pro Met Thr Phe Leu Ile Thr Ser Thr Asn As 35 40 45	rg
Thr Asn Arg Ser Ala Cys Leu Asp Leu Thr Ser Ser Asp Glu Leu As 50 55 60	sn
Thr Ile Lys Trp Tyr Asn Leu Ile Leu Thr Ala Ser Thr Phe Cys Le 65 70 75 8	eu ĉ
Pro Leu Val Ile Val Thr Leu Cys Tyr Thr Thr Ile Ile His Thr Le 85 90 95	e·u
Thr His Gly Leu Gln Thr Asp Ser Cys Leu Lys Gln Lys Ala Arg A. 100 105 110	rg
Leu Thr Ile Leu Leu Leu Leu Ala Phe Tyr Val Cys Phe Leu Pro Pi 115 120 125	he
His lie Leu Arg Val Ile Gln Asp Arg Ile Ser Ala Cys Phe Gln Ser 130 135 140	e, r
Val Val Pro Leu Arg Ile Arg Ser Met Lys Leu Thr Ser Phe Leu A. 145 150 155	.sp 60
His Tyr Ala Ala Leu Asn Thr Phe Gly Asn Leu Leu Leu Tyr Val V 165 170 175	āl
Val Ser Asp Asn Phe Gln Gln Ala Val Cys Ser Thr Val Arg Cys L 180 185 190	ýS
<pre><210: 61 <211: 872 <212: DNA <215: H.Sapiens</pre>	
$<\!400\%-61$ gggadggete gtagacacae taaccetace etttetgttt etteeteate titeet	ttcc 60
atotattet catggiotec tgtotgtotc tototototc coctofffor ototoc	taga 120
tettheteat eccetecatt tetgtgteaa teteaateea titatategg tggeea	cttt 180
totalctett tgttetatet eretetetet erettteeea etttgtetet geaege	ctgt 240
iğiginitte igeetgiete teletigese teatetetet giotetetet igeset	cate 300
tetergtote totgtgtotg tgtotococc gotcattocc atttgcaggt geastg	
aggadaacto atggageeee ceegggeeea tegagtaceg gaetggetga eeeset	
ttggdagtag cooctgacod toagtalqyd daadactadd ggagagddig aggagg	itgag 480

Page 36

nggenetetg tecceacegt eegeateage thatytgaag etgytaetge tgggaetgat	540									
'atg'gogtg agostggogg gtaasgooat ottgtoootg etggtgotsa aggagsgggo	€00									
netgoacaag gotoettaet acttoetget ggacetgtge etggeogatg geatacgete	660									
tgoogtotgo ticoccitty tgotggoite tgtgogocac ggotottoat ggacciteag	720									
tgcauteage tgcaagattg tggcotttat ggcogtgctc ttttgcttoc atgcggcott	780									
Hatgotgito typatcageg toaccegeta catggecate geocaccaec gettetaege	840									
caagngeatg acastotgga catgogogge tg	872									
-0.10: 62 -0.11: 143 -0.12: PRT -0.13: H.Sapiens										
Met Ala Asn Thr Thr Gly Glu Pro Glu Glu Val Ser Gly Ala Leu Ser 1 5 10 15										
Pro Pro Ser Ala Ser Ala Tyr Val Lys Leu Val Leu Leu Gly Leu Ile 20 25 30										
Met Cys Val Ser Leu Ala Gly Asn Ala Ile Leu Ser Leu Leu Val Leu 35 40 45										
Lys Glu Arg Ala Leu His Lys Ala Pro Tyr Tyr Phe Leu Leu Asp Leu 50 55 60										
Cys Lei Ala Asp Gly Ile Arg Ser Ala Val Cys Phe Pro Phe Val Leu 55 70 75 80										
Ala Mer Val Arg His Gly Ser Ser Trp Thr Phe Ser Ala Leu Ser Cys 85 90 95										
Lys Ile Val Ala Phe Met Ala Val Leu Phe Cys Phe His Ala Ala Phe 100 105 110										
Most Leu Phe Cys Ile Ser Val Thr Arg Tyr Met Ala Ile Ala His His 115 120 125										
Arg Phe Tyr Ala Lys Arg Met Thr Leu Trp Thr Cys Ala Ala Glu 130 135 140										
0210 + 63 0211 + 962 0212 + DNA 0213 + H.Sapiens										
$1400 \cdot -63$ Alaquitget glacigaact attgaatgga acttggaaat aaagteeett ecaadataac	60									

Page 37

tättottoaa dagagagtaa taggtaaatg tittagaagt gagaggacto aaattgodaa —— 120

tgatttactc ttttattt	tt ceteetaggt	ttctgggata	agtatgtgca	aataaaaaaat	180
aaacatgaga aggaactg	ta acctgattat	ggatttggga	aaaagataaa	tcaacacaca	240
aagggaaaag taaactga	tt gasayeeste	aggaatgitg	cccttttgcc	acaatataat	300
taatatttee tgtgtgaa	aa acsactggtc	aaatgatgte	cgtgetteec	tgtacagttt	360
autggtgete ataattet	ga coacactegt	tggcaatstg	atagttattg	tttctatatc	4170
acacttcaaa caacttca	ta odocaacaaa	ttggctcatt	cattccatgg	ccactgtgga	4811
ctttcttctg gggtgtct	gg tcatgcctta	cagtatgjtg	agatetgetg	ageactgttg	540
gtattttgga gaagtett	ct gtaaaattca	cacaagcacc	gacattatge	tgagetsage	600
ctccattttc catttgtc	tt teateteeat	tgaccgctac	tatgctgtgt	gtyatczact	61.11
gagatataaa gesaagat	ga atatotiggt	tatttgt jtg	atgatettea	ttagttggag	71.0
tgtoootgot gtttttgo	at ttggaatgat	ctttctgjag	ctaaacttca	aaggogotga	780
agagatatat tabaaaba	tg ttcactgcag	aggaggtige	tctgtcttct	ttagcaaaat	840
adetggggta etgacett	ta tgacttottt	ttatatacct	ggatctatta	tgttatgtgt	9(11)
chattacaga atatatot	ta togotaanga	acagycasga	ttaattagtg	atgscaatca	950
ga					962

#210: 64 #2212: 238 #212: PRT

+12131 H.Sapiens

:400:- 64

Arg Glu Lys Thr Asp Gln Pro Ser Gly Met Met Pro Phe Cys His Asn

He He Asn He Ser Cys Val Lys Asn Asn Trp Ser Asn Asp Val Arg

A.a Jer Leu Tyr Ser Leu Met Val Leu Ilé Ile Leu Thr Thr Leu Val

Gly Asn Leu Ile Val Ile Val Ser Ile Ser His Phe Lys Gln Leu His

Thr Pro Thr Ash Trp Leu Ile His Sor Met Ala Thr Val Asp Phe Leu 65

Leu Gly Cys Leu Val Met Pro Tyr Ser Mct Val Arg Ser Ala Glu His

Cys Trp Tyr Phe Gly Glu Val Phe Cys Lys Ile His Thr Ser Thr Asp $100 \,$ $\,$ $105 \,$ $\,$ $110 \,$

Ile Met Le		Ala Ser	Ile Ph	e His Le	u Ser Ph 12		e Ser Il	e
Asp Arg Ty 130	r Tyr Ala	Val Cys 135	Asp Pr	o Leu Ar	g Tyr Ly 140	's Ala	Lys Me	t
Asn Ile Le 145	u Val Ile	Cys Val 150	Met Il	e Phe Il 15		p Ser	Val Pr	
Aia Val Ph	e Ala Phe 165		Ile Ph	e Leu Gl 170	u Leu As	n Phe	Lys Gl 175	y
Ala Glu Gl	u Ile Tyr 180	Tyr Lys	His Va. 18		s Arg Gl	y Gly 190		r
Val Phe Phe 19		Ile Ser	Gly Va. 200	l Leu Th	r Phe Me 20		Ser Ph	e
Tyr lle Pro	o Gly Ser	Ile Met 215	Leu Cy	s Val Ty	r Tyr Ar 220	g Ile	Tyr Le	ບ
Ile Ala Ly. 225	s Glu Gln	Ala Arg 230	Leu Ile	e Ser As 23		n Gln	ı	
-0210 - 65 -0211 - 1010 -0212 - DNA -0213 - H.S.	8 apiens							
100 65								
nacautooog	ggtggaac	ct gggca	tgtat a	ttttgatt	g tittat	gcat	acticota	gtg 60
aagaaccaat	gtattgat	ca gatag	aagca a	gatactca	g acttac	tttc	totytag	ctc 120
stgctttta	ttattcct	gg ttgga	ttgca c	cactacto	a gittet	attt	tataata	etg 180
uttataaaac	atgggagg	ga aataa	ctttg t	attgyttt	t tatgga	taat	trattat	gtg 240
tectagacte	tggccttg	tc aaaag	aagga c	gtaagaag	g cacgat	gtat	tatactt	jgg 300
aatgatagaa	gagactga	cc tggta	tttcc a	cccggaag	a gggaaa	ggat	tttaact	aca 360
nataraggaa	tocagoag	at ggcat	cagag a	acactata	a annaga	наасд	atttgca	aca 410
jecanototo	ttocaaaa	ca attcc	ttact t	atgtagta	t gcaage	jeggt	tttttga	atg 480
gaacagaaca	tagtaata	ta ggaaa	acada a	tgatgaga	a aageca	пусаа	gttbaba	oct 540
gttggggaaa	agcacact	tt taaca	totca g	gegtaaaa	g tcaaca	igtaa	aattast	gtg 600
gtacaggttg	agtatoco	tt accca	aaatg t	ttgaaacc	a gaaatg	jt tit t	ggailttc	gga 660
nttoggaata	tttadada	tt cataa	tgata t	atottgga	a atggtt	acca	agtotaa	aca 720
maaa itttat	ttatgttt	ca tatac	acctt a	tacacata	g tetgaa	agta	attitgt	aca 780
∋tat•ttaaa	taattttq	gg catga	aacaa a	gtttgcat	a cattga	acca	toaqaca	gca 840
aaagitteag				gcatcatg				

960

1018

attttagago atttcaaatt ttggatttto aaattacaaa tgottaacct gtacttagat gttaaatana gtgoctotto dadgggdadt ttdaggaagd atlotittat ataagedd <210> 66 327 <211> ·:212:-PRT ·:213:-H.Sapiens ::100: 66 Tyr lle Lys Glu Cys Phe Leu Lys Val Pro Val Glu Glu Ala Leu Tyr Leu Thr Ser Lys Tyr Arg Leu Ser Ile Cys Asn Leu Lys Ile Glm Asn Leu Lys Cys Ser Lys Ile Trp Asn Phe Leu Ser Ile Asn Met Met Pro Gin Val Glu Asn Ser Thr Pro Glu Ala Phe Ala Val Trp Phe Asn Val CVs Lys Deu Cys Phe Met Pro Lys Ile Ile Asn Ile Val Gln Asn Tvr Phe Gln Thr Met Cys lle Arg Cys Ile Asn Ile Asn Lys Phe Cys Val Thr Trp 5lu Pro Phe Pro Arg Tyr Ile Ile Met Asn Val lle Phe Arg 105 Asn Pro Lys Ser Lys Thr Phe Leu Val Ser Asn Ile Leu Gly Lys Giv Tyr Ser Thr Cys Thr Thr Val Ile Leu Leu Leu Thr Phe Thr Pro Glu 135 Met Leu Lys Val Cys Phe Ser Pro Thr Gly Val Asn Leu Leu Ala Phe Leu Ile Ile Val Phe Scr Tyr Ile Thr Met Phe Cys Ser Ile Gln Lys Thr Ala Leu Gln Thr Thr Glu Val Arg Asn Cys Phe Gly Arg Glu Val Ala Val Ala Asn Arg Phe Phe Phe Ile Val Phe Ser Asp Ala Ile Cys Erp Ile Pro Val Phe Val Val Lys Ile Leu Ser Leu Phe Arg Val Glu lle Pro Gly Gln Ser Leu Leu Ser Phe Pro Ser Ile Ile His Arg Ala Phe Leu Arg Pro Ser Phe Asp Lys Ala Arg Val Asp Thr Ile Ile His

Lys Asn Gln Tyr Lys Val Ile Ser Leu Pro Cys Pne Ile Ile Ser Ile 270

Ile Lys Lys Leu Ser Ser Gly Ala Ile Gln Pro Gly Ile Ile Lys Ser 285

Arg Ser Tyr Arg Glu Thr Lys Ser Glu Tyr Leu Ala Ser Ile Ala Arg 290

His Trp Phe Phe Thr Arg Ser Met His Lys Thr Ile Lys Ile Tyr Met

Pro Arg Phe His Pro Gly Leu 325

<210: 67

GC11> 1251

<212> DNA

33133 H.Sapiens

(400) 67

actancatgg aagetgacet gggtgeeact ggeeacagge eeegcacaga gettgatgat 60 gaggaetest accessagg tggctgggas acggtettes tggtggeset getgetestt 120 quigotiques de caatiggett gatiggegtigg etgqcogget coraggeoog quatiggaget 180 queacgegte tggegetget extgeteage etggecetet etgaettett gtteetggea 2:0 quagraggest topagatest agagatesqq cataggggas actaggeoget ggggasaget 300 in stigeoget totactactt cotatgggge gtgtoctact esteoggest stiestigetg 366 greguectea geotogacog etgeotgoty gegetgtgee cacactygta coetggycae .; subsecaption quotiqueest staggitutge quotigitator aparticitage bacactette 320 adoptionet goetagtist becogaget geogtetigt ggradgaget ggscatétige 540 stagactict aggacagoga agagetated etgaggatge tagaggteet aggaggette 600 staggetttee teetactagt eatetaggeae attateagge aaggeaeagg stateaggaee turcacoque aacageagee ogeageetge oggggetteg eergtgtgge baggaeeatt 7110 chatcagest augusteet gaggetgeed taccagetag coragetget stacctaged 780 rtice: grupe aegintacte iggetacetg etetgggagg centggteta streegaetac 846 organization to advangety coloageces theoretyce testagecay typequete 900 regacectge tgegeteegt getetegtee ttegeggeag et@tetgega ggageggeeg 460 quitagettea egeccactga gecacagade cagetagatt etgagggtee aactetgeca 1026 1080 quiquequatiqui dagaqqoona qtoabaqatig igaticotigtigq (becaqeetca iggtigaaueec

Fage 41

acactecage caegategga teccacaget cagecacage tgaaccetae ggeecageca 1140
cagteggate ceacageeca gecacagetg aaccteatgg eccagecaca gteagattet 1200
gtggeecage caeaggeaga caetaaegte cagaeceetg caeetgetge c 1251

<2105 68 <2115 417 <2125 PRT

:213: H.Sapiens

4001 68

Thr Thr Met Glu Ala Asp Leu Gly Ala Thr Gly His Arg Prc Arg Thr 1 10 15

Glu Leu Asp Asp Glu Asp Ser Tyr Pro Gln Gly Gly Trp Asp Thr Val $20 \hspace{1cm} 25 \hspace{1cm} 30$

Pho Leu Val Ala Leu Leu Leu Gly Leu Pro Ala Asn Gly Leu Met 35 40 45

Ala Trp Leu Ala Gly Ser Gln Ala Arg His Gly Ala Gly Thr Arg Leu 50 60

Ala Leu Leu Leu Ser Leu Ala Leu Ser Asp Phe Leu Phe Leu Ala 45 70 75 80

Ala Ala Phe Gl
n Ile Leu Glu Ile Arg His Gly Gly His Trp Pro35
 30 95

Leu Gly Thr Ala Ala Cys Arg Phe Tyr Tyr Phe Leu Trp Gly Val Ser 100 105 110

Tyr Ser Ser Gly Leu Phe Leu Leu Ala Ala Leu Ser Leu Asp Arg Cys 115 120 125

Leu Leu Ala Leu Cys Pro His Trp Tyr Pro Gly His Arg Pro Val Arg i30 135 140

Leu Pro Leu Trp Val Cys Ala Gly Val Trp Val Leu Ala Thr Leu Phe 145 150 150 160

her Val Pro Trp Leu Val Phe Pro Glu Ala Ala Val Trp Trp Tyr Asp 165 170 175

theu Val Ile Cys Leu Asp Phe Trp Asp Ser Glu Glu Leu Ser Leu Arg 180 185 190

Met Leu Blu Val Leu Gly Gly Phe Leu Pro Phe Leu Geu Leu Leu Val 195 200 205

Cys His Val Leu Thr Gln Ala Thr Ala Cys Arg Thr Cys His Arg Gln 210 215 220

GIn Gln Pro Ala Ala Cys Arg Gly Phe Ala Arg Val Ala Arg Thr 11e 225 230 235 240

Leu	Ser	Ala	Tyr	Val 245	Val	Leu	Arg	Leu	Pro 250	Tyr	Gln	Leu	Ala	Gin 255	Leu	
Leu	Tyr	Leu	Ala 260	Phe	Leu	Trp	Asp	Val 265	Tyr	Ser	Gly	Туг	Leu 270	Leu	Trp	
Glu	Ala	Leu 275	Val	Tyr	Ser	Asp	Tyr 280	Leu	Ile	Leu	Leu	Asn 285	Ser	Суѕ	Leu	
Ser	Pro 290	Phe	Leu	Cys	Leu	Met 295	Ala	Ser	Ala	Asp	Leu 300	Arg	Thr	Leu	Leu	
Arg 305	Ser	Val	Leu	Ser	Ser 310	Phe	Ala	Ala	Ala	Leu 315	Cys	Glu	Glu	Arg	Pro 320	
Gly	Ser	Phe	Thr	Pro 325	Thr	Glu	Pro	Gln	Thr 330	Gln	Leu	Asp	Ser	Glu 335	Gly	
Pro	Thr	Leu	Pro 340	Glu	Pro	Met	Ala	Glu 345	Ala	Gln	Ser	Gln	Met 350	Asp	Pro	
Val	Ala	Gln 355	Pro	Gln	Val	Asn	Pro 360	Thr	Leu	Gln	Pro	Arg 365	Ser	Asp	Pro	
Thr	Ala 370	Gln	Pro	Gln	Leu	Asn 375	Pro	Thr	Ala	Gln	Pro 380	Gln	Ser	Asp	Pro	
Thr 385	Ala	Gln	Pro	Gln	Leu 390	Asn	Leu	Met	Ala	Gln 395	Pro	Gln	Ser	Asp	Ser 400	
Val	Ala	Gln	Pro	Gln 405	Ala	Asp	Thr	Asn	Val 410	Gln	Thr	Pro	Ala	Pro 415	Ala	
Ala																
-1210 -1211 -1212 -1213	.> 6 ∷≻ D	59 559 NA I.Sap	oiens	;												
-1400 taca		9 etg a	igcat	gata	ig gc	tcca	itcaç	, cac	caaq	jca-:	tged	etgte	ca t	cate	tggcc	60
date	tagt	ac c	egatig	rddad	c ac	ccca	caca	cct	gtca	igca	gtag	gtgtg	jtc (etget	ctggg	120
acat	gtcc	ct ç	jctgo	agag	ic at	acto	gaat	gga	itgtt	ctg	tggc	tto	etg (ctaç	gtggtg	180
otga	ittet	gt t	tggt	gtga	a ac	atca	igatt	tca	atcac	agt	caca	itggo	etg a	attt	tttat	240
gtgt	ggtt	ct d	etgeg	ngto	c aç	lacad	gttc	tgc	tggt	cag	gato	cttt	gt (ggato	ecgga	300
agat	gccc	ett ç	јасса	ugget	g ta	cato	jacca	too	etget	cag	agto	jetge	ito i	tect	cetet	360
qtga	idatg	icc c	otttg	ngcat	t da	igtga	itted	: tat	tttt	otg	gato	cacq	jtg (jattt	gtcac	420
gtto	gtct	ag t	ttod	attt	t co	etgto	cact	ctt	aaca	igca	gtgd	caa:	rcc (catta	atttac	480
ntat	tcat	gg ç	getec	ettta	ig ge	agct	tcaa	a aac		iaga age		ctac	gat (ggtta	ctecag	540

600 659

agggetetge aggaeaegee tgaggtggaa gaaggeagat ggeggettte tgaggaaaee														
etggagetgt cutgaageag attggggeea tgaggaagag eetetgeeet gteagteag														
<pre> (210: 70 (211: 213 (212: PRT (213: H.Sapiens)</pre>														
4.400: 7€														
Tyr Arg Bro Glu His Ala Gly Leu His Gln His Gln Ala Leu Pro Val 1 5 10 15														
His Pro Val Ala His Leu Val Pro Leu Pro Pro Pro His Thr Pro Val 20 25 30														
Ser Ser Arg Val Ser Cys Ser Gly Pro Cys Pro Cys Cys Arg Ala Ser														
Trp Asn Gly Cys Ser Val Ala Ser Cys Leu Val Val Leu Ile Leu Phe 50 60														
Gly Val Lys His Gln Ile Ser Ser Gln Ser His Gly Phe Phe Tyr Val $\frac{1}{65}$														
Trp Fhe Ser Ala Gly Pro Ala Arg Phe Cys Trp Ser Gly Ser Phe Val 85 90 95														
Asp Iro Gly Ang Cys Pro Pro Gly Cys Thr Pro Ser Cys Ser Glu Cys 100 100 100														
Trp Fer Ser Ser Ser Val Thr Cys Pro Leu Ala Phe Ser Asp Ser Tyr 115 120 125														
Phe Ser Gly Ser Thr Trp Ile Cys His Val Arg Leu Val Scr Ile Phe 130 135 140														
Lieu Ser Thr Leu Asn Ser Ser Ala Asn Pro Ile Ile Tvr Phe Phe Met 145 150 155 160														
Gly Mer Phe Arg Gln Leu Gln Ash Arg Lys Thr Leu Leu Val Leu Gln 165 170 175														
Arg Ala Leu Gln Asp Thr Pro Glu Val Glu Glu Gly Arg Trp Arg Leu 180 185 190														
Ser Glu Glu Thr Leu Glu Leu Ser Ser Arg Leu Gly Pro Gly Arg Ala 195 200 205														
Ser Ala Leu Ser Val 710														
1210 - 71 1211 - 559 1212 - DNA 1213 - H.Sapiens														

<400> 71	caggccgcag	aagagaagag	gaggacggtg	aggaggatga	acccadadaa	60
gooologgggt	gggggccgcc	gggggccccg	Crectar, coge	agcagcagca	caaggeeggs	120
cocacacatg	gtgcaacaca	gcagagocag	cagcaccgct	gccaccagcc	acagogtoog	180
gcacuagtgg	eggetggget	ccccgaagaa	ctgggtgcag	gcgccgctga	gcagcaggtg	240
cagcagcagg	cagagggccc	aggtgagggc	gcacadacag	gtggtcaggt	ggcgtgggcg	300
qcggcacgag	taccaggctg	ggaagagggc	ggccaggcac	tgctccacgc	tgacggcogc	360
caggagacto	aggcccacga	tgtagcagaa	gaagegeage	gttgccaggc	tggtctgsac	420
цаадпееддід	aagtccagcc	ggccttgcag	caagtegggg	acqatggcca	ccatgtgjca	480
gccaaggaag	atgagatoog	cgcaggccac	gtccaggagg	tagatggcga	aagggtttct	540
qtagacattg	gagetgage					559

H210H 72 H211H 211

-212: PRT

-213: H Sapiens

- 300: 72

Leu Ser Ser Asn Val Tyr Arg Asn Pro Phe Ala Ile Tyr Leu Leu Asp \pm 10 15

Val Ala Cys Ala Asp Leu Ile Phe Leu Gly Cys His Met Val Ala Ile 20 \$25\$

Val Fro Asp Leu Leu Gln Gly Arg Leu Asp Phe Pro Gly Phe Val Gln 35° 40° 45°

Thr Ser Leu Ala Thr Leu Arg Phe Phe Cys Tyr Ile Val Gly Leu Ser 50 60

Leu Leu Ala Ala Val Ser Val Glu Gln Cys Leu Ala Ala Leu Phe Pro 6^6 70 75 80

Ala Trp Tyr Ser Cys Arg Arg Pro Arg His Leu Thr Thr Cys Val Cys 85 90

Alà Leo Thr Trp Ala Leo Cys Leo Leo Leo His Leo Thr Thr Cys Val 100 105 110

Cys Ala Léu Thr Trp Ala Leu Cys Leu Leu His Leu Leu Leu Ser $115 \,$ $120 \,$ $125 \,$

Gly Ala Cys Thr Leu Leu Leu Ser Gly Ala Cys Thr Gln Phe Phe Gly 130 140

Glu Pro Ser Arg His Leu Cys Arg Thr Leu Trp Leu Val Ala Ala Val 145 150 160

Leu Leu Ala Leu Leu Cys Cys Thr Met Cys Gl 165 170	y Ala Ser Leu Met Leu 175
Leu Leu Arg Val Glu Arg Gly Pro Gln Arg Pr 180 195	ro Pro Pro Arg Gly Phe 190
Pro Gly Leu Ile Leu Leu Thr Val Leu Leu Ph 195 200	ne Ser Ser Ala Ala Cys 205
Leu Arg His	
:210 - 73 :221 - 1008 :212 - DNA :213 - H.Sapiens	
:400 73 htgg:atcat officealt tggagtgato bffgctgtc	to tygocicoot cateattyct 60
actancacae tagtggctgt ggetgtgttg etgttgate	oc acaagaatga tygtgtcagt 120
ototiqotica cottigaatot ggotigtiggot gacacettig	ga ttggtgtgge catetotgge — 180
otacicacag accagotote cagooottot oggoocaca	ac agaagaccot gtgcagcotg 240
ogganggeat tigicacite ofpogoaget geotetgic	ed toacggteat gotgatizach 300
ttigalaggt acctigedat daagsagedd tioogdiad	et tgaagatcat gagtgggtte 366
gtiggwogggg cotgeattge egggetgtgg ttagtgtet	t acoteantgg offoctocea 400
micgyaatoo obatgttooa goagaotgoo tabaaaggg	ge agtgeagett etttgetgta 485
titoscocco actiogigot gascalotec igogitiggo	of fottoecage patgetooth 541
titgrattet tetäetgega eatgeteaag attgeetee	ca tghacagoca geagattega — 600
akganggaad atgoaggago batgyotiga ggttatoga	at occomingad toopagegas 665
thoawageto toogracigt glotightobe altigggage	t tigototate oligiaccocc 700
tico: tatoa otggoatigt geaggingles incompagag	gt gtoaceteta cetagrigetig 780
gaacogtace tgtggetget eggegtggge aacteeetg	go toaacecact catotatges 849
tattugcaga aggaygtgog actgbagetb taccabatg	gg cootaggagt gaagaaggtg 900
bloatoteat tectestett teleteggee aggaattgt	g geocagagag geocagggaa 960
agtt-otgto acatogtoac tatotocago toagagtit	g atggetaa 1008

12:10 · 74 12:11 · 335 12:12 · PRT 12:13 · H.Sapiens

Page 46

WO 01/36473

<400> 74

Met Glu Ser Ser Phe Ser Phe Gly Val Ile Leu Ala Val Leu Ala Ser Leu Ile Ile Ala Thr Asn Thr Leu Val Ala Val Ala Val Leu Leu Leu Ile His Lys Asn Asp Gly Val Ser Leu Cys Phe Thr Leu Asn Leu Ala Val Ala Asp Thr Leu Ile Gly Val Ala Ile Ser Gly Leu Leu Thr Asp Gln Leu Ser Ser Pro Ser Arg Pro Thr Gln Lys Thr Leu Cys Ser Leu Arg Met Ala Phe Val Thr Ser Ser Ala Ala Ala Ser Val Leu Thr Val Met Leu Ile Thr Phe Asp Arg Tyr Leu Ala Ile Lys Gln Pro Phe Arg Tyr Leu Lys Ile Met Ser Gly Phe Val Ala Gly Ala Cys Ile Ala Gly Leu Trp Leu Val Ser Tyr Leu Ile Gly Phe Leu Pro Leu Gly Ile Pro Met Phe Gln Gln Thr Ala Tyr Lys Gly Gln Cys Ser Phe Phe Ala Val 145 150 155 160 Phe His Pro His Phe Val Leu Thr Leu Ser Cys Val Gly Phe Phe Pro Ala Met Leu Leu Phe Val Phe Phe Tyr Cys Asp Met Leu Lys Ile Ala 185 Ser Met His Ser Gln Gln Ile Arg Lys Met Glu His Ala Gly Ala Met Ala Gly Gly Tyr Arg Ser Pro Arg Thr Pro Ser Asp Phe Lys Ala Leu Arg Thr Val Ser Val Leu Ile Gly Ser Phe Ala Leu Ser Trp Thr Pro Pho Leu Ile Thr Gly Ile Val Gln Val Ala Cys Gln Glu Cys His Leu Tyr Leu Val Leu Glu Arg Tyr Leu Trp Leu Leu Gly Val Gly Asn Ser Leu Leu Asn Pro Leu Ile Tyr Ala Tyr Trp Gln Lys Glu Val Arg Leu Gln Leu Tyr His Met Ala Leu Gly Val Lys Lys Val Leu Thr Ser Phe

Leu Leu Phe Leu Ser Ala Arg Asn Cys Gly Pro Glu Arg Pro Arg Glu 305 310 315 320

Sor Sor Cys His Ile Val Thr Ile Ser Ser Ser Glu Phe Asp Gly 325 330 335

1210: 75

:211: 2137 ::212: DNA

.213: H.Sapiens

·:400:- 75

aactggaagg geageegtet geogeogang aacasettet caageasttt gagtgaceae 157 agottiquaaq etqqtqqctq qoocoocqaq teocqqqqcte tqaqqcacqq boqteqaott 1.20 aageqttqsa teetqttaee tqqaqaeset etqaqetete acetqetaet tetqeeqetq 130 cttotqcaca qaqooqqqq qaqqacocot ocaqqatqca qqtcccqaac aqcaccqqcc 240 eggacaacge gacgetgeag atgetgegga accepgegat egeggtggee etgeoegtgg 3190 360 tytactoget ggtggeggeg gebageatee egggbaaeet offetetetg togggtgetgt qooggogoat qgggoodaga tooboqtoqg toatottoat gatoaacetq aqoqtbacqq 4.10 apetgatijet ggeoagogtij tigootities aaatotasta seatigoias bijosaseast 2000 gagtattogg ggtgotgott tgoaabgtgg tgabbgtggd bittitabgba aacatgtatt 5:10 scaqeatest caccatques tigtatouged tiggagagett cotiggigists etigtaceage BUIL tradotrosa godotogodo odcogtogot acgoquitogo ogcquetigoa guqaoctogo Sec. O. 720 tgotgotoot gaoogoodig toocogotigg ogogosooga totoacctae boggiqoacg 786 dectoggest categoodge tregsegted teaagtggsd gatgeteled agegtggses 840 tgtgggdegt gttoctette accatottba tectgetgtt cotdatoobg thogtgatda sequagetty thacaegges accatected agetyttysy saeggaggag gegeaegges 96.0 aggaagaaga gaagagagaa gaagagaataa caqagataat attqataaca titqataacat 960 1000 gettequese caacaactte gtgeteetgg egeacategt gageegesig thetaeggea 1080 agagetaeta ecaegigiae aageteaege igigieteag eigeeteaae aacigietigg acceptitigt thattactit gegiceeggg aastocaget gegeetgegg gwatattigg gatgoogoog ggtgddbaga gabaccotgg acacgngoog ogagagooto tholocycca 1290 1.2 mg qqaccacqto oqtgoqotoc gaggooggtg oqcabbobga agggatgqag gqagbbabba 1320 ggeneggest coagaggeag gagagtgtgt totgagtese gggggegoag ettggagage ogggogogea gottggagga todaggggog datggagagg odaoggtgod agaggttdag 1330 1440 ggagaacago tqoqttqoto ocaggoactq cagaggooog gtggggaligg gtotocaggo Page 48

tttattcctc	ccaggcactg	cagaggcasc	ggtgaggaag	ggtctccagg	cttcactcag	1500
ggtagagaaa	caagcaaagc	ccagcagcjc	acagggtgct	tgttatcctg	cagagggtgc	1560
atatgaatat	ctgtgtcagg	ggacagettg	tgtcaccacg	cccggctaat	ttttgtattt	1620
ttttagtag	agctgggctg	tcacccccga	geteettaga	cactcctcac	acctglocat	1680
accegaggat	ggatattcaa	ccagececae	cgcctacccg	actcggtttc	tggatateet	1740
ctgtyggcga	actgcgagcc	ccattcccag	ctcttctccc	tgctgacatc	gtcccttagc	1800
acacctgtcc	atacccgagg	atggatattc	aaccagcccc	accgcctacc	cgacteggtt	1860
tetggatate	ctctgtgggc	gaactgcgag	coccattocc	agetettete	cctgctgaca	1920
negtecetta	gttgtggttc	tggccttctc	cattotocto	caggggttct	ggtotoogta	1980
geceggtgea	cgccgaaatt	tctgtttatt	tcactcaggg	gcactgtggt	tgctgtggtt	2040
ggaattette	tttcagagga	gcgcctgggg	ctcctgcaag	teagetacte	teegtgeeea	2100
ottoocotca	cacacacacc	cccctcgtgc	cgaattc			2137

Met Gln Val Pro Asn Ser Thr Gly Pro Asp Asn Ala Thr Leu Gln Met

Leu Arg Asn Pro Ala Ile Ala Val Ala Leu Pro Val Val Tyr Ser Leu 20 25 30

Val Ala Ala Val Ser Ile Pro Gly Asn Leu Phe Ser Leu Trp Val Leu

Cys Arg Arg Met Gly Pro Arg Ser Pro Ser Val Ile Phe Met Ile Asn 50 60

Leu Ser Val Thr Asp Leu Met Leu Ala Ser Val Leu Pro Phe Gln fle

Tyr Tyr His Cys Asn Arg His His Trp Val Phe Gly Val Leu Leu Cys

Asn Val Val Thr Val Ala Phe Tyr Ala Asn Met Tyr Ser Ser Ile Leu

Thr Met Thr Cys Ile Ser Val Glu Arg Phe Leu Gly Val Leu Tyr Pro

Leu Ser Ser Lys Arg Trp Arg Arg Arg Tyr Ala Val Ala Ala Cys 130 135 140

^{+310&}gt; 76 +311> 359 +312> PRT +313> H.Sapiens

^{-:400:- 75}

Ala 145	Gly	Thr	Trp	Leu	Leu 150	Leu	Leu	Thr	Ala	Leu 155	Ser	Pro	Leu	Ala	Arg 160	
Thr	Asp	Leu	Thr	Tyr 165	Pro	Val	His	Ala	Leu 170	Gly	Ile	Ile	Thr	Cys 175	Phe	
Asp	Val	Leu	Lys 180	Trp	Thr	Met	Leu	Pro 185	Ser	Val	Ala	Met	Trp 190	Ala	Val	
Phe	Leu	Phe 195	Thr	Ile	Phe	lle	Leu 200	Leu	Phe	Leu	lle	Pro 205	Phe	Val	Ile	
Thr	Val 210	Ala	Cys	Tyr	Thr	Ala 215	Thr	Ile	Leu	Lys	Leu 220	Leu	Arg	Thr	Glu	
Glu J25	Ala	His	Gly	Arg	Glu 230	Gln	Arg	Arg	Arg	Ala 235	Val	ej A	Leu	Ala	Ala 240	
Val	Val	Leu	Leu	Ala 245	Phe	Val	Thr	Cys	Phe 250	Ala	Pro	Asn	Asn	Phe 255	Val	
Leu	I.eu	Ala	His 260	lle	Val	Ser	Arg	Leu 265	Phe	Tyr	Gly	Lys	Ser 270	Туг	Tyr	
His	\'al	Tyr 275	Lys	Leu	Thr	Leu	Cys 280	Leu	Ser	Cys	Leu	Asn 285	Asn	Cys	Leu	
Asp	Pro 290	Phe	Val	Tyr	Туr	Phe 295	Ala	Ser	Arg	Glu	Phe 300	Gln	Leu	Arg	Leu	
Arg 305	Glu	Tyr	Leu	Gly	Cys 310	Arg	Arg	Val	Pro	Arg 315	Asp	Thr	Leu	Asp	Thr 320	
Arg	Arg	Glu	Ser	Leu 325	Phe	Ser	Ala	Arg	Thr 330	Thr	Ser	Val	Arg	Ser 335	Glu	
Ala	С1у	Ala	His 340	Pro	Glu	Gly	Met	Glu 345	Gly	Ala	Thr	Arg	Pro 350	Gly	Leu	
Gln	Arg	Gln 355	Glu	Ser	Val	Phe										
+(210 -(210 -(210 -(210	1. · · · 1 2 · · · I	77 1197 DNA H.Sa _l	oien:	5												
-:40(atq		7.7 Bara (ggete	getgo	og ge	cagge	egees	g gto	gagoo	jägg	tcat	tagta	act (geati	tacaac	60
_				-					-						gaegee	120
gtg	gtgt	gad t	ggc	ggtgl	Lg Cq	gasti	tcato	e gto	gctaq	gaga	atct	tage	egt (gttgi	ttggtg	180
cte	ggac	god a	accc	geget	tt c	palego	ctcc	s atq	gttc	ctgc	teet	t giggi	cag	ecto	acgttg	240
teg	gatet	tge 1	tggc:	agge	ga a	geet	aege	e ger	caac	atcc	tact	tgtc	jgq ·	ghogi	etcacg	300

ctgaaa	ctgt	acceageget	ctggttcgca	cgggaggjag	gogtottogt	ggcactcact	360
gagtica	gtgc	t gagectect	ggccatcgcg	ctggagejca	gootcaccat	ggogogoagg	420
gggccc	gcgc	cogtotocag	tegggggege	acgctgg:ga	tggcagccgs	ggcctggggc	43.)
gtgtcg	ctgc	toctoggget	cctgccagcg	ctgggctgga	attgcctggg	togootggac	540
gettget	tcca	ctgtcttgcc	gctctacgcc	aaggcctacg	tgctcttctg	egtgetegee	600
ttcgtg	ggca	tectggeege	tatctgtgca	ctctacgcgc	gcatctactg	ccaggtacgc	660
уссаар	gcgc	ggcgcctgcc	ggcacggccc	gggactgcgg	ggaccacctc	gaccegggeg	720
ogtogoa	aagc	agagatagat	ggccttgctg	cgcacgctca	gegtggtget	cetggeettt	780
itggcat	tgtt	ggggeeeect	cttcctgctg	ctgttgctcg	acgtggcgtg	eceggegege	840
acctgto	ectg	tactcctgca	ggeogatece	ttcctgggac	tiggocatggo	caactcactt	900
ctgaaco	ccca	teatetacae	gctcaccaac	ogogacotgo	gocaegeget	cetgegeetg	960
utotgot	tgcg	gacgccactc	ctgcggcaga	gacccgagtg	gotoccagoa	gteggegage	1020
gaggatq	gagg	cttccggggg	cctgcgccgc	tgcctgcccc	cgggccttga	tigggagette	1080
agoggot	cgg	agegeteate	geeccagege	gacgggctgg	acaccagogg	ctccacagge	1140
ageodoc	ggtg	cacccacage	cgcccggact	ctggtatcag	ancoggotgo	agactga	1197

-210 - 78

-211 398

·212 · PRT

· 213 · H.Sapiens

-400: 78

Met Glu Ser Gly Leu Leu Arg Pro Ala Pro Val Ser Glu Val I.e Val I $10\,$

The Ile Val Leu Glu Asn Leu Ala Val Leu Leu Val Leu Gly Arg His

Fro Arg Phe His Ala Pro Met Phe Leu Leu Leu Gly Ser Leu Thr Leu $\epsilon.5$ 70 75 80

Cer Asp Leu Leu Ala Gly Ala Ala Tyr Ala Ala Asn Ile Leu Leu Ser

Gly Pro Leu Thr Leu Lys Leu Ser Pro Ala Leu Trp Phe Ala Arg Glu 100 105 110

G1 y	Gly	Val 115	Phe	Val	Ala	Leu	Thr 120	Ala	Ser	Val	Leu	Ser 125	Leu	Leu	Ala
Ile	Ala 130	Leu	Glu	Arg	Ser	Leu 135	Thr	Met	Ala	Arg	Arg 140	Gly	Pro	Ala	Pro
Val 145	Ser	Ser	Arg	Gly	Arg 150	Thr	Leu	Ala	Met	Ala 155	Ala	Ala	Ala	Trp	Gly 160
Val	Ser	L∈u	Leu	Leu 165	Gly	Leu	Leu	Pro	Ala :70	Leu	Gly	Trp	Asn	Cys 175	Leu
Gly	Arq	L∈u	Asp 180	Ala	Cys	Ser	Thir	Val 185	Leu	Pro	Leu	Tyr	Ala 190	Lys	Ala
Tyr	Va]	Leu 195	Phe	Cys	Val	Leu	Ala 200	Phe	Val	Gly	Il∈	Leu 205	Ala	Ala	Ile
Cys	Ala 210	Leu	Tyr	Ala	Arg	Ile 215	Tyr	Cys	Gln	Val	Arg 220	Ala	Asn	Ala	Arg
Arg 225	Leu	Pro	Ala	Arg	Pro 230	Gly	Thr	Ala	Gly	Thr 235	Thr	Ser	Thr	Arg	Ala 240
Arg	Arg	Lys	Pro	Arg 245	Ser	Leu	Ala	Leu	Leu 250	Arg	Thr	Leu	Ser	Val 255	Val
Leu	Leu	Ala	Phe 260	Val	Ala	Cys	Trp	GLy 265	Pro	Leu	Phe	Leu	Leu 270	Leu	Leu
≟⊕u	Asp	Val 275	Ala	Суѕ	Pro	Ala	Arg 280	Thr	Cys	Pro	Val	Leu 235	Leu	Gln	Ala
Азр	Pro 290	Phe	Leu	Gly	Leu	Ala 295	Met	A.l.ā	Asn	Ser	Leu 300	L∙∍u	Asn	Pro	Ile
:le 305	Tyr	Thr	Leu	Thr	Asn 310	Arg	Asp	Leu	Arg	His 315	Ala	Leu	Leu	Arg	Leu 320
Val	Су.з	Cys	Gly	Arg 325	His	Ser	Cys	Gly	Arg 330	Asp	Pro	Ser	Gly	Ser 335	Gln
Gln	Ser	Ala	Ser 340	Ala	Ala	Glu	Ala	Ser 345	Gly	Gly	Leu	Arg	Arg 350	Cys	Leu
Pro	Pro	Gly 355	Leu	Азр	Gly	Ser	Phe 360	Ser	Gly	Ser	Glu	Arg 365	Ser	Ser	Pro
Gln	Arg 370	Asp	Gly	Leu	Asp	Thr 375	Ser	Gly	Ser	Thr	Gly 380	Ser	Pro	Gly	Ala
Pro 385	Thr	Ala	Ala	Arg	Thr 390	Leu	Val	Ser	Glu	Pro 395	Ala	Ala	Asp		
<21 <21 <21 <21	1:- 2:-	79 1041 DNA H.Sa	pien	s											

						<400→ 79	
60	gatgeegeeg	teteccaggt	ggggacacca	ccgcatcgag	ggtcgtgctg	atgtacaacg	
120	gtgtggttts	gggtcgccct	ctaggcaatg	gctgggggca	tggcctttgt	ctgctcattg	
180	ggccgtggct	ttttcaatit	actgtttacc	gaageccage	tgaagacctg	tgottocaca	
240	acgtagacac	attacctcag	cggacagact	cctgcctttt	ttatgatctg	gatttcctcc	
300	gaacagggee	cgttggccat	gjgctcttca	ctgccgagtg	gggacattcc	ngggettttg	
360	ggtccacccc	atttcaaagt	goggadaggt	ggtggtggct	tgttccttac	ggagcatcg	
420	caccctgtgg	gcatcgtctg	gtggcggctg	ctccacccgg	tgaacactat	caccacgcgg	
480	cgtgcaagag	accatctctg	tigotggaga	agtgtatctt	tootgggaac	goodlygtda	
540	tgacatcatg	atggctggca	gagteggeea	cttcatcatg	cstgtgagag	abggoogtet	
600	caagattgtt	tttgctcctt	atcatcttat	geceetegge	agttctttat	ttccagctgg	
650	ggcgacccgg	ggatgaagaa	agacaggete	gcagetggee	ggcggaggca	tggageetga	
720	gtctgctaga	tgcccagcgt	adatgotado	tgtgttcatc	tggtggcaat	ttcafcatgg	
780	tggggeeetg	cctctgtcca	geetgegate	gccetcgagt	tetggaeggt	ctctatttcc	
840	gtattatttt	atecectggt	ageatgetgg	ctacatgaac	teagetteae	cheataacce	
900	gaaacccaag	tctgcagtct	aagctcaaaa	attotacaac	cettteccaa	.caagecect	
960	ceteggtege	caatttcgaa	gaagagatgc	acaaaggccg	actcaaaaac	cugedaggae	
1010	atgggatccc	ctgatgggca	caaagccagt	aaatagtttc	tcagtgLggc	aggaqttgca	
1041				a	agtggcactg	dasa:tgttg	

+0010 + 80 +0011 + 346 +0012 + PRT +0013 + H.Sapiens

-1400 · 80

Met Tyr Asn Gly Ser Cys Cys Arg Ile Glu Gly Asp Thr Ile Ser Gin

Val Met Pro Pro Leu Leu Ile Val Ala Phe Val Leu Gly Ala Leu Gly

Asn Gly Val Ala Leu Cys Gly Phe Cys Phe His Met Lys Thr Trp Lys

Pro Ser Thr Val Tyr Leu Phe Asn Leu Ala Val Ala Asp Phe Leu Leu

Met Ile Cys Leu Pro Phe Arg Thr Asp Tyr Tyr Leu Arg Arg Arg His 65 70 75 80

PCT/US00/31581 WO 01/36473

Trp	Ala	Phe	Gly	Asp 85	Ile	Pro	Cys	Arg	Val 90	Gly	Leu	Phe	Thr	Leu 95	Ala
Met	Asn	Arg	Ala 100	G.y	Ser	11e	Val	Phe 105	Leu	Thr	Va1	Val	Ala 110	Ala	Asp
Arg	Tyr	Phe 115	Lys	Val	٧al	His	Pro 120	His	His	Ala	Val	Asn 125	Thr	Ile	Ser
Thr	Arg 130	Val	Ala	Ala	Gly	Ile 135	Val	Cys	Thr	Leu	Trp 140	A.a	Leu	Val	.1€
Leu 145	Gly	Thr	"al	Tyr	Leu 150	Leu	Leu	Glu	Asn	His 155	Lou	Cys	Val	Gln	Glu 160
Thr	Ala	Val	Ser	Cys 165	Glu.	Ser	Phe	Ile	Met 170	Glu	Ser	A_a	Asn	G1y 175	Trp
His	Asp	Ile	Met 180	Pi:e	Glr	Leu	Glu	Phe 185	Phe	Met	Pro	Leu	Gly 190	11e	Ile
Leu	Phe	Cys 195	Ser	Phe	Lya	Tle	Val 200	Trp	Ser	Leu	Arg	Arg 205	Arg	Glr	Gln
Leu	Ala 210	Arg	Gin	Ala	Arg	Met 215	Lys	Lys	Ala	Thr	Arg 2.10	Phe	Ile	Met	Val
Уа1 225	Ala	Ile	Val	Phe	11e 230	Thr	Cys	Tyr	Leu	Pro 235	Ser	Val	Ser	Ala	Ara 240
Leu	Tyr	Phe	Leu	Trp 245	Thr	٧al	Pro	Ser	Ser 250	Ala	Cys	Азр	Pro	Ser 255	Va:
His	Gly	Ala	Leu 260	H_s	Ilm	Thr	Leu	Ser 265	Phe	Thr	Tyr	Net	Asn 270	Ŝeı	Met
Leu	Asp	Pro 2.5	Leu	Veil	Tyr	Tyr	Phe 280	Ser	Ser	Pro	Ser	Phe 235	Pro	Lys	Ph∈
Tyr	Asn 290	Lys	Leu	Lys	Ile	Cys 295	Ser	Leu	Lys	Pro	Lys 300	Gln	Pro	Gly	ніг
Ser 345	Lys	Thr	Gln	Arg	Prο 310	Glu	Glu	Met	Pro	11e 315	S∻r	Aan	Leu	Gly	Arc 320
Ang	Ser	Cys	Ile	Ser 325	Vаl	Ala	Asn	Ser	Phe 330	Gln	Ser	Gln	Ser	Asp 335	G15
Gln	Trp	Asp	Pro 340	His	Il:	Val	Glu	Trp 345	His						
<21: <21: <21: <21	1:- ; 2:- i	81 2525 DNA H.Sa	pien.	s											

<213: H.Sapiens

<400 - 81

caagaatgac aggtgactte scaagtatge etggecacaa tacetecagg aatteetett 60 Page 54

gegatectat	agtgacaccc	casttaatca	gcctctactt	catagtgctt	attggcgggc	120
tggtgggtgt	catttccatt	cttttcctcc	tggtgaaaat	gaacacccgg	tcagtgacca	180
ccatggcggt	cattaacttg	gtggtggtcc	acagcgtttt	tetgetgaea	gtgccatttc	240
gcttgaccta	cctcatcaag	aajacttgga	tgtttgggct	gcccttctgc	aaatttgtga	300
gtgccatgct	gcacatccac	atgtacctca	cgttcctatt	ctatgtggtg	atcctggtca	360
ccagatacct	catcttcttc	aagtgcaaag	acaaagtgga	attctacaga	aaactgcatg	4.30
ctgtggctgc	cagtgctggc	atgtggacgc	tggtgattgt	cattgtggta	cccctggttg	480
tctcccggta	tggaatccat	gaggaataca	atgaggagca	ctgttttaaa	tttcacaaag	540
agettgetta	cacatatgtg	aaaatcatca	actatatgat	agtcattttt	gtcatagecg	600
ttgctgtgat	tctgttggtc	ttocaggtot	tcatcattat	gttgatggtg	cagaagetae	1511()
gesactettt	actateceae	caggagttct	gggctcagct	gaaaaaccta	ttttttatag	7.10
gggtcatect	tgtttgtttc	cttccctacc	agttetttag	gatctattac	ttgaatgttg	780
tgacgcattc	caatgcctgt	aacagsaagg	ttgcatttta	taacgaaatc	ttottgagtg	840
taacagcaat	tagotgotat	gatttgcttc	tetttgtett	tgggggaage	cattggttta	900
agcaaaagat	aattggctta	tggaattgtg	ttttgtgccg	ttagccacaa	actacagtat	Sect
tcatatttgc	tteetttata	ttgygaataa	aaatgggtat	aggggaggta	agaatggtat	1000
ttdattactt	gatcaaaacc	atgeettgat	gtacccaaaa	caaaaggact	ataaaatgca	1090
agagccetea	ttgtagtcct	tatgggatcc	ctcccatctc	tgagtgatgg	ccgtacaaag	1140
accagtgttg	ttgaatccac	ctggagttgc	aatattacat	tattttccag	tacagaatgt	1700
ctgtgtggcc	catgaaagca	acataggttt	taagagtttt	agagtttcat	tageteatte	1260
taagtteete	tgtttgaage	atggtctctt	aggttttgga	ctgaactcag	acctttagtt	1320
cttttcatcc	cacttcacct	taggtaagta	aattetggee	accacccagc	tocaaagaca	1380
cadactotec	ttegetaace	aggttagatg	teccatteat	ctcatgeect	gataaaaact	1440
gataagggga	gagaatagtt	aaaaattttt	ctagggtatc	ataactctgg	taggaagtica	$1 f \otimes \emptyset$
toigtotaga	aatcaagaga	aaaagaacgt	gtggcctcct	gttataacaa	gggtttetag	1000
atttgtcctg	tgaaaggtog	tttaaggact	tggggatcaa	cttcctcaat	tateaceaat	1+4.0
tgcactgttg	ctccaaaaat	catttaäaag	cttactggac	atarctacat	aatggtgaaa	1680
otgtaattta	gagactatee	ctgactaatg	tgctggtagg	cattaaaatg	agttcccaag	1740
ggaagtgatt	aaaatttttt	totottotgt	tttttgagag	aatttetaga	tgtcctgggc	1800

Page 55

cacagttaat	taagattttt	aggggggaca	gaaagttata	ctgaaatctt	tagagetece	1860
tteegeegtt	aaaattatat	atatatatat	ttaaattata	ccttaagttc	tggggtacat	1920
gtgcagaatg	tgcaggtttg	ttacataggt	atacacgtgc	catggtggtt	tgcggcacct	1980
gtcaacccat	ctacattagg	tatttctcct	aatgetetee	ctcccctage	cocceaccc	2049
tggacaggcc	ccattgtgtg	atgttcccct	ccctgtgtcc	atgtgtttt:	attgttcaac	2100
teceaettet	aagtgagaac	atgcggtgtt	tggttttctg	ttcctgtgtt	agittgctga	2160
gaatgatggt	ttccaggtta	aaattatata	tttttaaata	aatgaaaact	gtgtttttaa	2200
aagaggactt	ttgagaagta	tatagaaaaa	ccattaattt	agactctgtg	agattaggtt	2280
gcatgaagaa	ggttttctga	atatttgaag	agtgjataaa	taaatgtcc:	ccaaagcaat	2340
aaaatcataa	tootttaaaa	tataggaaaa	ātaastaatg	ggaactaggs	ttaatactcg	24(10
ggatgaaata	atotgtacaa	caaactecca	tgacacatgt	ttacctatgt	aacaaacctg	2460
cacatgtacc	cctgaactta	aaataaaatt	taaajtataa	taataaaata	atatggattt	2520
tottt						2525

^{∹210 - 82}

<400 - 82

Met Thr Gly Asp Phe Pro Ser Met Pro Gly His Asn Thr Ser Arg Asn

Ser Ser Cys Asp Pro Ile Val Thr Pro His Leu lle Ser Leu Tyr Pho

The Val Lou The Gly Cly Leu Val Gly Val The Ser The Leu Phe Leu

Leu Val Lys Met Asn Thr Arg Ser Val Thr Thr Met Ala Val Ile Asn

Leu Val Val Val His Ser Val Phe Leu Leu Thr Val Pro Phe Arg Leu

Thr Tyr Leu Ile Lys Lys Thr Trp Met Phe Gly Leu Pro Phe Cys Lys

Phe Val Ser Ala Met Leu Eis Ile His Met Tyr Leu Thr Phe Leu Phe 105

Tyr Val Val Ile Leu Val Thr Arg Tyr Leu Ile Phe Phe Lys Cys Lys 120

Asp Lys Val Glu Phe Tyr Arg Lys Lou His Ala Val Ala Ala Ser Ala Page 56

	130					135					140				
Gly 145	Met	Trp	Thr	Leu	Val 150	Ile	Val	Ile	Val	Val 155	Pro	Leu	Val	Val	Ser 160
Arg	Tyr	Gly	Ile	His 165	Glu	Glu	Tyr	Asn	Glu 170	Glu	His	Cys	Phe	Lys 175	Phe
Hıs	Lys	Glu	Leu 180	Ala	Tyr	Thr	Tyr	Val 185	Lys	Ile	Ile	Asn	Tyr 190	Met	Ile
Val	Ile	Phe 195	Val	Ile	Ala	Val	Ala 200	Val	Ile	Leu	Leu	Val 205	Phe	Gln	Val
Phe	Ile 210	Ile	Met	Leu	Met	Val 215	Gln	Lys	Leu	Arg	His 220	Ser	Leu	Leu	Ser
His 225	Gln	Glu	Phe	Trp	Ala 230	Gln	Leu	Lys	Asn	Leu 235	Phe	Phe	Ile	Gly	Val 240
Ile	Leu	7al	Cys	Phe 245	Leu	Pro	Tyr	Gln	Phe 250	Phe	Arg	Ile	Tyr	Tyr 255	Leu
Asn	Val	Val	Thr 260	His	Ser	Asn	Ala	Cys 265	Asn	Ser	Lys	Val	Ala 270	Phe	Tyr
Asn	Glu	11e 275	Phe	Leu	Ser	Val	Thr 280	Ala	Ile	Ser	Cys	Tyr 285	Asp	Leu	Leu
Leu	Phe 290	Val	Phe	Gly	Gly	Ser 295	His	Trp	Phe	Lys	Gln 300	Lys	Ile	Ile	Gly
Lou 305	.Trp	Asn	Cys	Va1	Leu 310	Cys	Arg								
-:210 -:211 -:212 -:21.	! ·	33 1125 DNA H.Sap	piens	5											
-:400 gcae		33 cac 1	tgaaa	aatca	ag ga	nacaa	atcet	: gta	attt	tttg	tgat	aato	caa (caago	gacaaa
act:	tata	cat a	atgta	aaata	ia ca	agcgt	:tato	g ago	cagea	aatt	cate	ecat	get (jgt.gc	getgtg
cago	etgto	got a	acgco	gaacq	gt ga	atg	ggtad	c tg:	igtga	aaaa	tcc	ectto	ctc (jeeg	ggatco
ogge	g+gai	tte 1	tgtad	cata	gt gt	ttg	gatit	gg	ggcto	gtgc	tggd	ctgt	gtt	tggaa	aacctc
stge	gt.gat	tga :	ttta	aatco	at c	catt	саас	g dag	gctgo	cact	ctc	egaco	caa	tttt	ctcgtt
geet	tate	tgg (ectg	egete	ga ti	tctt	ggt	g ggt	tgtga	actg	tgat	gee	ett -	cagea	atggtc
agga	acgg:	taja (agago	etget	ig gt	atti	tggg	g age	gagti	tttt	gtad	sttt	cca ·	cacot	tgctgt

jatgtggcal titgttacto tietetetit caettgtget teateteeat egacaggtae

attgeggtta etgacecect ggtetatect accaagitea eegtatetgt gleaggaatt

Page 57

420

480

540

tgcatcageg	tgtcstggat	catgacaata	atgtacageg	gtgctgtgtt	ctacacaggt	600
gtctatgacg	atgggctgga	ggaattatct	gatgccctaa	actgtatagg	ajgttgtcag	660
accgttgtaa	atcaasactg	ggtgttgaca	gattttctat	sottotttat	abotacettt	720
attatgataa	ttotgtatgg	taacatattt	cttgtggcta	gacgacaggc	gaaaaagata	780
gaaaatactg	gtagcaagac	agaatcatcc	tcagagagtt	эсаладссад	agtggccagg	840
agagagaa	aagcagctaa	aaccctgggg	gtcacagtgg	tagcatttat	gatttcatgg	900
ttaccatata	gcattgattc	attaattgat	gcctttatgg	gctttataac	ceetgeetgt	960
atttatgaga	tttgctgttg	gtgtgcttat	tataactcag	scatgaatco	tttgatttat	1020
gotthattt	acccatggtt	taggaaagca	ataaaagtta	ttgtaactgg	teaggtttta	1080
aagaacagtt	cagcaaccat	gaatttgttt	totgaacata	tataa		1125

<f.10:- 84

4211: 345

RT H.Sapiens

-(400: 84

Mot Ser Ser Asn Ser Jer Leu Leu Val Ala Val Gin Leu Cys Tyr Ala 10 15

Akh Mal Ash Gly Ser Cys Val Lys Ile Pro Phe Ser Pro Gly Ser Arg 20 - 25 - 30

Val The Leu Tyr The Val Phe Giy Phe Gly Ala Val Leu Ala Val Phe 35 - 40

Gly Asn Leu Leu Wal Met Ile Ser Ile Leu His Phe Lys Gln Leu His '0 55 60

Ser Pro Thr Ash Phe Leu Val Ala Ser Leu Ala Cys Ala Asp Phe Leu 55 70 75 80

Val Cly Val Thr Val Met Pro Phe Ser Met Val Arg Thr Val Glu Ser 95 99

Gys Trp Tyr Phe Gly Arg Ser Phe Gys Thr Phe His Thr Gys Gys Asp $100 \,$ $105 \,$ $110 \,$

 $761~\mathrm{Ala}$ Phe Cys Tyr Ser Ser Leu Phe His Leu Cys Phe Ile Ser Ile 115 120 125

Asp Arg Tyr Ile Ala Val Thr Asp Pro Leu Val Tyr Pro Thr Lys Phe

Thr Mal Ser Val Ser Gly He Cys The Ser Val Ser Trp He Leu Pro 145 150 155 160

Leu Met Tyr Ser Gly Ala Val Phe Tyr Thr Gly Val Tyr Asp Asp Gly Page 58

WO 01/36473	PCT/US00/3158

	165		170		175
Leu 31u Glu Le 1-	eu Ser Asp 30	Ala Leu Asr 185		Gly Gly Cys 190	Gln Thr
Val Val Asn G. 195	in Asn Trp	Val Leu Thr 200	Asp Phe	Leu Ser Phe 205	Phe Ile
Pro Thr Phe I.	le Met Ile	Ile Leu Tyr 215	Gly Asn	Ile Phe Leu 220	Val Ala
Arg Arg Gln A. 225	La Lys Lys 230	Ile Glu Asr	Thr Gly 235	Ser Lys Thr	Glu Ser 240
Ser Ser Glu Se	er Tyr Lys 245	Ala Arg Val	Ala Arg 250	Arg Glu Arg	Lys Ala 255
Ala Lys Thr Le	eu Gly Val რმ	Thr Val Val 265		Met Ile Ser 270	Trp Leu
Pro Tyr Ser I 275	le Asp Ser	Leu Ile Asp 280	Ala Phe	Met Gly Phe 285	Ile Thr
Pro Ala Cys II .190	le Tyr Glu	Ile Cys Cys 295	Trp Cys	Ala Tyr Tyr 300	Asn Ser
Ala Met Asn Pi 305	ro Leu Ile 310	Tyr Ala Leu	ı Phe Tyr 315	Pro Trp Phe	Arg Lys 320
Ala ile Lys Va	al Tie Val 325	Thr Gly Glr	Val Leu 330	Lys Asn Ser	Ser Ala 335
Thr Met Asn Le	eu Phe Ser 40	Glu His Ile 345			
<pre><b10: 1020="" 85="" <b110:="" <b12:="" <b13:="" dna="" h.sapie<="" pre=""></b10:></pre>	ens				
4400: 85 accatgaatg age	ecactaga ct	atttagda aa	itacttota	attteccega t	tatgcaget 60
qotthtggaa ath	-	<u>-</u>	_	-	
tatggcatta to:	itootogt gg	gattteca go	gcaatgcag	tagtgatate o	cacttacatt 180
titicanaatya gad	octtggaa ga	gcagcacc at	cattatgc	tgaacetgge :	etgcacagat 240
ongoigtaic tga	accagost co	ectteetg at	tcactact	atgccagtgg c	gaaaactgg 300
atotitggag att	ttcatgtg ta	agtttate eq	geticaget	tocatttcaa k	ectgtatage 360
agoar octot too	ctcacstg tt	tcagcate tt	ccgctact	gtgtgatcat t	cacccaatg 420
agetqetttt eca	attcacaa aa	ictogatgt go	cagttgtag	cotgtgctgt (ggtgtggatc 480
atttbactgg tac	getgteat to	egatgace tt	cttgatca	catcaaccaa o	caggaddaad 540

agateageet	gtetegaeet	caccagttcg	gatgaactca	atactattaa	gtqqtacaac	600
ctgattttga	ctgcaagtac	tttatgaata	cccttggtga	tagtgacact	tigotataco	650
acgattatcc	acactitgac	ccatggantg	caaactgaca	gctgccttaa	gcagasagcs	720
cgaaggotaa	ccattetget	actecttyca	ttttacgtat	gtttttacc	cttccatatc	780
ttgagggtca	ttcaggatcg	aatctcaccc	tgctttcaat	cagttgttcc	attgagaatc	840
agatocatga	agettacate	gtttctadac	cattatgctg	ctctgaacac	otttggtaac	900
ctgttactat	atgtggtggt	cagegacaac	tttcagcagg	ctgtctgctc	aacagtgaga	960
tqcaaagtaa	gegggaacet	tgagcaagca	aagaaaatta	gttactcaaa	caaccottga	1020

⁻C2105 86

<4000 ≥ 86

Met Asn Glu Pro Leu Asp Tyr Leu Ala Asn Ala Ser Asp Phe Pro Asp 1 510

Tyr Ala Ala Ala Phe Gly Asn Cys Thr Asp Glu Asn Ile Pro Leu Lys 20 - 25 30

Met His Tyr Leu Pro Val Ile Tyr Gly Ile Ile Phe Leu Val Gly Phe 35 40 45

Pro Gly Asn Ala Val Val Ile Ser Thr Tyr Ile Phe Lys Met Arg Pro 50 55 60

Trp Lys Ser Ser Thr Ile Ile Mot Leu Ash Leu Ala Cys Thr Asp Leu 65 70 75 80

Leu Tyr Leu Thr Ser Leu Pro Phe Leu Ile His Tyr Tyr Ala Ser Gly 85 90 95

Glu Ann Trp Ile Phe Gly Asp Phe Met Cys Lys Phe Ile Arg Phe Ser 100 105 110

Phe His Phe Asn Leu Tyr Ser Sor Ile Leu Phe Leu Thr Cys Phe Ser 115 125

Ile Ph.e Arg Tyr Cys Val Ile Ile His Pro Met Ser Cys Phe Ser Ile

His Lys Thr Arg Cys Ala Val Val Ala Cys Ala Val Val Trp Ile Ile 145 155 160

Ser Leu Val Ala Val Ile Pro Met Thr Phe Leu Ile Thr Ser Ihr Asn 165 170 175

Arg Thr Ash Arg Ser Ala Cys Leu Asp Leu Thr Ser Ser Asp 3lu Leu 180 185 190

^{+211&}gt; 336

^{-:212&}gt; PRT

⁻⁰⁰¹³⁰ H.Sapiens

Asn	Thr	Ile 195	Lys	Trp	Tyr	Asn	Leu 200	Ile	Leu	Thr	Ala	Ser 205	Thr	Phe	Суѕ	
Leu	Pro 210	Leu	Val	Ile	Val	Thr 215	Leu	Cys	Tyr	Thr	Thr 220	lle	Ile	His	Thr	
Leu 225	Thr	His	Gly	Leu	Gln 230	Thr	Asp	Ser	Суѕ	Leu 235	Lys	Gln	Lys	Ala	Arg 240	
Arg	Leu	Thr	Ile	Leu 245	Leu	Leu	Leu	Ala	Phe 250	Tyr	Val	Cys	Phe	Leu 255	Pro	
Phe	His	Ile	Leu 260	Arg	Val	Ile	Gln	Asp 265	Arg	Ile	Ser	Ala	Cys 270	Phe	Gln	
Ser	Val	Val 275	Pro	Leu	Arg	Ile	Arg 280	Ser	Met	Lys	Leu	Thr 285	Ser	Phe	Leu	
Asp	His 290	Tyr	Ala	Ala	Leu	Asn 295	Thr	Phe	Gly	Asn	Leu 300	Leu	Leu	Tyr	Val	
Val 305	Val	Ser	Asp	Asn	Phe 310	Gln	Gln	Ala	Val	Cys 315	Ser	Thr	Val	Arg	Cys 320	
Lys	Val	Ser	Gly	Asn 325	Leu	Glu	Gln	Ala	Lys 330	Lys	Ile	Ser	Tyr	Ser 335	Asn	
-4210 -211 -:212 -:213	.> 1	87 .138 NA i.Sap	oiens	5												
.400):+ E	17														
нааа	atto	jet <u>e</u>	gtact	gaac	t at	tgaa	tgga	a act	t gga	aat	aaaq	gtaca	tt d	ccaaa	ataac	60
±a5t	.atta	aa c	cagac	gagta	ıa ta	nggta	iaatç	; ttt	taga	agt	gaga	iggac	ito a	aatt	gccaa	17.0
:gat	ttac	tc t	ttta	atttt	tco	et.cct	aggt	. ttc	rtggç	jata	agta	tgtc	jca a	ataa	aaaat	180
aac	atga	ıga a	iggaa	ctgt	аас	ectga	ittat	gga	tttç	gga	aaaa	igata	aaa t	caac	acaca	240
aagç	gaaa	ag t	aaac	etgat	t ga	icago	cata	: agg	aato	jatg	ccct	tttç	jed a	ıcaat	ataat	3(.0
r. aa t	attt	cc t	gtgt	gaaa	a ac	dact	ggto	aaa	itgat	gto	cgtg	gette	add t	gtac	agttt	36.0
uato	gtgc	tc a	taat	toto	a co	acac	tegt	t gg	gcaat	ctg	atag	ttat	tg t	ttet	atato	420
acac	ttca	aa c	caact	tcat	асс	ccaa	caaa	tte	gete	att	catt	ccat	.gg c	cact	gtgga	4 & 0
nttt	ctto	tg ç	ggtç	gtoto	g to	atgo	octta	a cag	gtatç	ytg	agat	.ctgc	etg a	igeac	etyttg	540
gtat	tttg	ıga ç	gaagt	cttc	t gt	aaaa	ttca	cae	aago	acc	gaca	ttat	iga t	gago	teage	60-0
ritod	attt	ta c	attt	gtct	t to	atot	.ccat	. tga	accgo	tac	tato	gctgt	igt ç	gtgat	ccact	660

gagatataaa gocaagatga atatottggt tatttgtgtg atgatottca ttagttggag

igtecotget gittitgeat tiggaatgat chitetggag etaaactica aaggegetga

Page 61

710

780

BN500x010 - Vv0 - 1136473A3

aga	gatai	tat 1	tacaa	aacat	tg t	tcact	tgcad	g ago	gaggt	ttga	toto	gtcti	tot :	ttago	caaaat
atc	tagga	gta	ctja	cctti	ta t	gacti	tatti	t tta	atata	acct	ggat	cta	tta	tgtta	atgtgt
cta	ttaca	aga a	atata	atctt	ta t	agat.	aaaga	a aça	agge	aaga	ttaa	attad	gtg .	atge	caatca
gaa	gete	caa a	attg	gatto	gg a	aatga	aaaa	a tgo	gaati	tca	caaa	адсаа	aag (aaago	gaaagc
tgt	gaaga	aca 1	tt-gg:	ggatt	tg to	gatgo	ggagt	t tit	ccta	aata	tgal	ggto	gcc (cttt	ctttat
etg	tacaç	gtc a	atgga	accet	tt ti	totto	cacta	a cat	tati	cca	acta	a-stt1	tga a	atgai	tgta
-(21) -(21) -(21) -(21)	1 · 2	88 296 PRT H.Sap	piens	3											
:40)	0 8	38													
Met l	Met	Pro	Phe	Cys 5	His	Asn	Ile	Ile	Asn 10	Ile	Ser	Cys	Val	Lys 15	Asn
Asn	Trp	Ser	Asn 20	Asp	Val	Arg	Ala	Ser 25	Leu	Tyr	Ser	Leu	Met 30	Val	L∈·u
Me	fle	Leu 35	Thr	Thr	Leu	Val	Gly 40	Asn	Leu	Ile	Val	Ile 45	Val	Ser	Ile
Ser	His 50	Phe	Lys	Gln	Leu	His 55	Thr	Pro	Thr	Asn	Trp 60	Leu	;le	His	Ser
Met 65	Ala	Thr	Val	Asr.	Phe 70	Leu	Leu	Gly	Cys	Leu 75	Val	Met	Pro	Tyr	Ser 80
Met	7al	Arg	Ser	Ala 85	Glu	His	Cys	Trp	Туг 90	Phe	Gly	Glu	Val	Phe 95	Cys
Lys	Ile	His	Thr 100	Ser	Thr	Asp	Ile	Met 105	Leu	Ser	Ser	Ala	Jer 110	Ile	Phe
Нιѕ	Leu	Ser 115	Phe	I16-	Ser	Ile	Asp 120	Arg	Tyr	Tyr	Ala	7al 125	Cys	Asp	Pro
Lou	Arg 130	Tyr	Lys	Ala	Lys	Met 135	Asn	Ile	Leu	Val	11e 140	Cys	Val	Met	Ile
Phe	He	Ser	Trp	Ser	Val		Ala	Val		Ala		G1y	Met	Ile	Phe 160

Leu Glu Leu Asn Phe Lys Gly Ala Glu Glu Ile Tyr Tyr Lys His Val

His Dys Arg Gly Gly Cys Ser Val Phe Phe Ser Lys Ile Ser Gly Val

Leu Thr Phe Met Thr Ser Phe Tyr Ile Pro Gly Ser Ile Met Leu Cys

Page 62

Val Tyr Tyr Arg Ile Tyr Leu Ile Ala Lys Glu Gln Ala Arg Leu Ile Ser Asp Ala Asn Gln Lys Leu Gln Ile Gly Leu Glu Met Lys Asn Gly Le Ser Gln Ser Lys Glu Arg Lys Ala Val Lys Thr Leu Gly Ile Val Met Gly Val Phe Leu Ile Cys Trp Cys Pro Phe Phe Ile Cys Thr Val

Met Asp Pro Phe Leu His Tyr Ile Ile Pro Pro Thr Leu Asn Asp Ala

Arg Gly Ser Arg Ala Asn Ser Ala 290

€.100 89

d:11/2 1023
d:12/2 DNA
d:13/2 H.Sapiens

·:00 · 89

qqaatqatqc cottttqcca caatataatt aatatttcct qtqtaaaaa caactqqtca 60 autgatgtod gtgottocot gtacagttta atggtgotoa taattotgad dadactogtt 120 quoautotga tagttattgt ttotatatea cactteaaac aactteatac cecaacaaat 180 tage catte attecatege castetegas tttetteteg getetetegt cateestad 240 antatogtoa gatotootoa goactottoo tattttoogag aagtottoto taaaattoac 300 acaageaceg acattatget gagetbagee tecattitiee attiquetti catetbeatt 3460 quecquiant atgrigtiqtiq tqatocactq aqatataaaq ccaaqatqaa tatqttqqtt 420 atttgtgtga tgatottcat tagttggagt gtocotgotg tttttgcatt tggaatgato 480 tttotggago taaacttoaa aggogotgaa gagatatatt acaaacatgt toactgoaga 540 gnaggttgot otgtottott tagoaaaata totggggtad tgaootttat gaottotttt 600 tatatacotg gatetattat gttatgtgte tattacagaa tatatettat egetaaagaa FIFILL caqqcaagat taattagtga tgocaatcag aagotocaaa ttggattgga aatgaaaaat 720 qqaatttcac aaagcaaaqa aaqqaaagct qtqaaqacat tqqqqqattqt qatqqqaqtt 780 ttoctaatat gotgqtgood tttotttato tqtacaqtea tggaccettt tottcactae 840 attattocae etaettigaa igaigiattig attiggittig getaettigaa etelaeatti 900 aatocaatgg tttatgcatt tttctstoct tggtttagaa aagcactgaa gatgatgctg 960 tttggtaaaa ttttccaaaa ägattcatcc aggtgtaaat tatttttgga attgagttca 1020

tag 1023

<210> 90 <211> 339 <212> PRT

-213: H.Sapiens

. 400:- 90

Het Met Pro Phe Cys His Asn Ile Ile Asn Ile Ser Cys Val Lys Asn I 10 10 15

The The Leu Thr Thr Leu Val Gly Asn Leu Ile Val The Val Ser Ile 35 - 40

Ger His Phe Lys Gln Lcu His Thr Pro Thr Asn Trp Leu lle His Ser 50 60

Met Ala Thr Val Asp Phe Leu Leu Gly Cys Leu Val Met Pro Tyr Ser 65 76 75 80

Met Val Arg Ser Ala Glu His Cys Trp Tyr Phe Gly Glu Val Phe Cys ± 5 90 95

Lys lie His Thr Ser Thr Asp Ile Met Leu Ser Ser Ala Ser Ile Phe $100 \,$.10

His Leu Ser Pho Ile Ser Ile Asp Arg Tyr Tyr Ala Val Cys Asp Pro 115 120 125

Leu Arg Tyr Lys Ala Lys Met Asn Ile Leu Val Ile Cys Val Met Ile 130 140

Phe lle Ser Trp Ser Val Pro Ala Val Phe Ala Phe Gly Met Ile Phe 145 150 160

Leu Glu Leu Asn Phe Lys Gly Ala Glu Glu Ile Tyr Tyr Lys His Val165 170 175

His Cys Arg Gly Sly Cys Ser Val Phe Phe Ser Lys ILe Ser Gly Va. 180 190

Leu Thr Phe Met Thr Ser Phe Tyr Ile Pro Gly Ser Ile Met Leu Cys 195 200 205

Tal fyr Tyr Arg lle Tyr Leu Ile Ala Lys Glu Sin Ala Arg Leu Ile

Ser Asp Ala Ash Gln Lys Leu Gln He Gly Leu Glu Met Lys Ash Gly

The Ser Gin Ser Lys Glu Arg Lys Ala Val Lys Thr Leu 31y Ile Val 245 250 255

Met Gly Val Phe Leu Ile Cys Trp Cys Pro Phe Phe Ile Cys Thr Val Page 64

Met Asp $\frac{1}{275}$ Phe Leu His Tyr $\frac{1}{280}$ Ile Pro Pro Thr Leu Asp Asp Val $\frac{1}{280}$ Leu $\frac{1}{280}$ Pro $\frac{1}{280}$ P

Gly Lys Ile Phe Gln Lys Asp Ser Ser Arg Cys Lys Leu Phe Leu Glu 325 330 335

Leu Ser Ser

H2105 91 H2117 1696 H2125 DNA

<213 H.Sapiens

·:400 · 91 stgtaaaqta qattqtatqa qqacteeatq aggteateea etteaagtee ttggeatagg €0 ataartacto aaaaqqtqat qacaatqqoq caqqqaqqqa tqqtqacttq octqqaqatq 120 cacaqeaceq tototoccat actoggicat teacaceate atigaticae caggeaceae 180 tecqtqteca qeaqqaetet qqqqaeeeca aatggacaet accatggaag etgacetggg 240 tgeenetgge caeaggeece geacagaget tgatgatgag gastectace cocaaggtgg 360 ongggacaeg gtetteetgg tygecetget geteentygg etgecageea atgggttgat 360 agedingety geogratices aggiological togagetyde acquitetyd cyclyctost 4200 deteageety decetetety actiettytt eetygeagea geggeettee agateetaga 4800 gationggout gggggacact ggccyctgyg gadagotgod tgccgottot actaettoot 540 Atgaggegte tectactest segmentatt cotyctyges gesetsages tegacogety 600 settestague etatacocae actigatacee tagacacego coagtoogoo taccoctota (5 + , 11 gatelgegee gqtqtetqqq tgetqgeese actettesge gtgeeetgge tgqtettees 7,10 780 equipment of the total transfer and the total transfer of the transfer and the transfer of the 840 geographic aggregation aggreeating aggreeating posttocated the the total 900 ctgccacqtg ctcacccaqq ccacaqcctq togcacctgc caccqccaac ayeaqcccqc aucottocog ggottogodo gtotagodag gaccattoto toagoctato togotoctoaq 960 getgeestae eagetggees agetgeteta estggeette etgtgggaeg totaetetgg 1020 1080 ctabetgete taggaggees taggtetacte egactaceta atectactea acagetgeet

cagocootto	ctctgcctca	tggccagtgc	cgasetecgg	accetgetge	gctccgtgct	1140
ctegteette	geggeagete	totgogagga	geggeeggge	agetteaege	ccactgagec	1200
acagacccag	ctagattetg	agggtccaac	totgocagag	cogatggcag	aggeceagte	1260
acagatggat	cctgtggccc	agcctcaggt	gaaccccaca	ctccagccac	gatoggatoc	1320
cacageteag	ccacagetga	accetaegge	ccagccacag	toggatecca	cagoccaged	1380
acagotgaso	ctcatggccc	agocacagto	agattotgtg	geccagedae	aggeagaeae	1440
taacgtccag	acccctgcac	ctgctgccag	ttotgtgccc	agtocotgtg	atgaagette	1500
occaacocca	tootograto	ctaccccagg	ggebettgag	gacccagcca	cacctcctgc	1560
ctctgaagga	gaaagcccca	geageacece	gccagaggcg	geceegggeg	caggccccac	1620
gtgagggtcc	aggaacacgc	aggeceacea	gagsagtgaa	agageccagg	gcagacagag	1680
gaaccagcca	gtcaga					1696

1210 - 92

<211: 505
<212 PRT</pre>

<213 H.Sapiens

-:100 93

Leu Ala Trp Arg Cys Thr Ala Pro Ser Leu Pro Tyr Ser Val Ile His l 10 15

Thr :le He Asp Ser Pro Gly Thr Thr Pro Cys Pro Ala Gly Leu Trp 20 25 30

Gly Pro Gin Met Asp Thr Thr Met Glu Ala Asp Leu Gly Ala Thr Gly 35 40 45

His Arg Pro Arg Thr Glu Leu Asp Asp Glu Asp Ser Tyr Pro Gln Gly

Gly Trp Asp Thr Val Phe Leu Val Ala Leu Leu Leu Leu Gly Leu Pro 65 70 80

Ala Ash Gly Leu Met Ala Trp Leu Ala Gly Ser Gin Ala Arg His Gly 95 95

Ala Gly Thr Arg Lei Ala Leu Leu Leu Lei Ser Leu Ala Leu Ser Asp

Pne Lei Pho Leu Alm Ala Ala Ala Phe Gln Ile Leu Glu Ile Arg His 115 120 125

Gly Gly His Trp Pro Leu Gly Thr Ala Ala Cys Arg Phe Tyr Tyr Phe 130 135 140

Leu Trp Gly Val Ser Tyr Ser Ser Gly Leu Phe Leu Leu Ala Ala Leu 145 150 150 160

Ser	Leu	Asp	Arg	Cys 165	Leu	Leu	Ala	Leu	Cys 170	Pro	His	Trp	Tyr	Pro 175	G1y
His	Arg	Pro	Val 180	Arg	Leu	Pro	Leu	Trp 185	Val	Cys	Ala	G17.	Val 190	Trp	Val
Leu	Ala	Thr 195	Leu	Phe	Ser	Val	Pro 200	Trp	Leu	Val	Phe	Pro 205	Glu	Ala	Ala
Val	Trp 110	Trp	Tyr	Asp	Leu	Val 215	Ile	Cys	Leu	Asp	Phe 220	Trp	Asp	Ser	Glu
31u 225	Leu	Ser	Leu	Arg	Met 230	Leu	Glu	Val	Leu	G1y 235	Gly	Phe	Leu	Pro	240
Leu	Leu	Leu	Leu	Val 245	Cys	His	Val	Leu	Thr 250	Gln	Ala	Thr	Ala	Cys 255	Arg
Thr	Cys	His	Arg 260	Glr.	Gln	Gln	Pro	Ala 265	Ala	Cys	Arg	Gly.	Phe 270	Ala	Arg
Val	Ыlа	Arg 275	Thr	Ile	L€:u	Ser	Ala 280	Tyr	Val	Val	Leu	Arg 285	Leu	Pro	Tyr
Gln	Leu 190	Ala	Gln	Leu	Leu	Tyr 295	Leu	Ala	Phe	Leu	Trp 300	Asp	Val	Tyr	Ser
31y 305	Tyr	Leu	Leu	Trp	Glu 310	Ala	Leu	Val	Tyr	Ser 315	Asp	Tyr	Leu	Ile	Leu 320
Leu	Asn	Ser	Cys	Leu 325	Ser	Pro	Ph∈	Leu	Cys 330	Leu	Met	Alá	Ser	Ala 335	Asp
Leu	Arg	Thr	Leu 340	Leu	Arg	Ser	Val	Leu 345	Ser	Ser	Phe	Alā	Ala 350	Alā	Leu
Cys	Glu	G1u 355	Arg	Pro	Gly	Ser	Phe 360	Thr	Pro	Thr	Glu	2ro 365	Gln	Thr	Gln
Leu	Asp 170	Ser	Glu	Gly	Pro	Thr 375	Leu	Pro	Glu	Pro	Met 380	Ala	Glu	Alâ	Gln
Ser 385	Gln	Met	Asp	Pro	Val 390	Ala	G1 rı	Pro	Gln	Val 395	Asn	Pro	Thr	Leu	G1n 400
Pro	Arg	Ser	Asp	Pro 405	Thr	Ala	Gln	Pro	Gln 410	Leu	Asn	Pro	Thr	Ala 415	Gln
Pro	Gla	Ser	Asp 420	Pro	Thr	Ala	Gln	Pro 425	G1n	Leu	Asn	Leu	Met 430	Ala	Gln
Pro	Gln	Ser 435	Asp	Ser	Val	Ala	Gln 440	Pro	Glr.	Ala	Asp	Thr 445	Asn	Val	Gln
	450					455					460		Asp		
Ser	Pro	Thr	Pro	Ser	Ser	Нis	Pro	Thr		Gly age		Leu	Glu	Asp	Pro

PCT/US00/31581

480 465 470 475 Ala Thr Pro Pro Ala Ser Glu Gly Glu Ser Pro Ser Ser Thr Pro Pro 485 195 Gru Ala Ala Pro Gly Ala Gly Pro Thr 93 <2100 1413 :211: .012. ANG (213) H.Sapiens 34000 93 atggacacta ccatggaage tgacetgggt gecactggce acaggeceeg cacagagett -50 quatquigagg actectacco colangitage taggacacag tettoctagit ageocetacta 120 proceedings typicagocia tyggittgatg gogtggotgg boggotocca ggocoggicat 120 ggagntggea egegtetgge getgeteetg eteageetgg eestetetga ettettgtte 240 stageageag eggeettesa gateetagag ateeggeatg ggggacaetg googetgagg 300 adagetgoot geogetteta etaetteeta tygggogtyt eetaeteete eggeetette Buch 4.0 otgotiggeng cootcagest egaledgetge etgetggege tgtgddcada etggtadddt ing success carteegest geseteting gtothering gtothering gtothering A < 0540 staticagog tgosotggst ggtattodos gaggatgoog tatggtggta ajasatggta atetqeetgg acttetgggs cagegaggag etgtegetga ggatgetgga ggteetgggg Er Hill اللبيون quetractiquatitacticat gotgeticate igeologitas telecological calcagodist 7.10 equancing acceptance gragecogies gentledging gettlegening tigtiggening 7200 accantetat careetatus queestqaqq etqeectace agetggeeca getgetetae orgeouted tgtgggaegt stactorgge taborgerer gggaggeest ggretaeres 840 gaetacetga testactesa caguigesto agusestico tetgestest ggosagigos 3 10 9.50 questioning bestgetign; sterritigets tegtesting egycagetet stigegaggag pggpogggea gottoaegee paptgageea pagapppage tagattotga gjjjtocaact 1020 ongonagago ogatggbaga ggbbbagtba bagatggatb btytggbbba gbbtbaggtig 1030 114:) Augumentac tecagedacy atogyatoco acagetbago capagetyaa bootaogyud dagbrabagt eggatedead ageodageda dagetgaade teatggeden gedadagtea 1200 gastutitigg eccagodada ggoagadast aacgtobaga obsctgoads tigotigodagt 126) 1320 totalgocea geocetatga tgaagettee coaaseseat estequates tacceeaggg 1380 acceptuagg acceagocae acctectuee tetqaaqqaq aaaqceccag caycaccecg

Page 68

WO 01/36473

ccagaggegg eccegggege aggececacg tga

1413

#:2105 #:211: #:212: #:213:	94 419 PRT H.Sapien	15										
·;400:-	94											
Met Asp 1	Thr Thr	Met Glu 5	Ala	Asp	Leu	Gly 10	Ala	Thr	Gly	His	Arg 15	Pro
Arg Thr	Glu Leu 20	Asp Asp	Glu	Asp	Ser 25	Tyr	Pro	Gln	Gly	Gly 30	Trp	Asp
Thr Val	Phe Leu 35	. Val Ala	Leu	Leu 40	Leu	Leu	Gly	Leu	Pro 45	Ala	Asn	Gly
Leu Met 50	: Ala Trp	Leu Ala	Gly 55	Ser	Gln	Ala	Arg	His 60	Gly	Ala	Gly	Thr
Arg Lev 65	ı Ala Leu	Leu Leu 70	Leu	Ser	Leu	Ala	Leu 75	Ser	Asp	Phe	Leu	Phe 80
Leu Ala	a Ala Ala	Ala Phe 85	Gln	I1e	Leu	Glu 90	Ile	Arg	His	Gly	Gly 95	His
Trp Pro	Leu Gly 100		Ala	Суѕ	Arg 105	Phe	Tyr	Tyr	Phe	Leu 110	Trp	Glγ
Val Ser	Tyr Ser 115	Ser Gly	Leu	Phe 120	Leu	Leu	Ala	Ala	Leu 125	Ser	Leu	Asp
Arg Cys 130	: Leu Leu)	Ala Leu	Cys 135	Pro	His	Trp	Tyr	Pro 140	Gly	His	Arg	Pro
Val Arg 145	Leu Pro	Leu Trp 150	Val	Cys	Ala	Gly	Val 155	Trp	Val	Leu	Ala	Thr 160
Len Phe	e Ser Val	Pro Trp 165	Leu	Vāl	Phe	Pro 170	Glu	Ala	Ala	Val	Trp 175	Trp
Tyr Asp	Leu Val 180	Ile Cys	Leu	Asp	Phe 185	Trp	Asp	Ser	Glu	Glu 190	Leu	Ser
Leu Arg	Met Leu 195	Glu Val	Leu	Gly 200	Gly	Phe	Leu	Pro	Phe 205	Leu	Leu	Leu
Leu Val 210	Cys His	Val Leu	Thr 215	Gln	Ala	Thr	Ala	Cys 220	Arg	Thr	Cys	His
Arg Gln 225	Gln Gln	Pro Ala 230	Ala	Cys	Arg	Gly	Phe 235	Ala	Arg	Val	Ala	Arg 240
Thr Ile	e Leu Ser	Ala Tyr 245	Val	Val	Leu	Arg 250	Leu	Pro	Tyr	Gln	Leu 255	Ala

PCT/US00/31581

WO 01/36473

Gln Leu Leu Tyr Leu Ala Phe Leu Trp Asp Val Tyr Ser Gly Tyr Leu 265 Leu Trp Glu Ala Leu Val Tyr Ser Asp Tyr Leu Ile Leu Leu Asn Ser Cys Leu Ser Pro Phe Leu Cys Leu Met Ala Ser Ala Asp Leu Arg Thr Leu Leu Arg Ser Val Leu Ser Ser Phe Ala Ala Ala Leu Cys Slu Glu Ard Fro Gly Ser Phe Thr Pro Thr Glu Pro Gln Thr Gln Leu Asp Ser Glu Gly Pro Thr Leu Pro Glu Pro Met Ala Glu Ala Gln Ser Gln Met Asp Fro Val Ala Gln Pro Gln Val Asm Pro Thr Leu Gln Pro Arg Ser Asp Fro Thr Ala Gin Pro Gln Leu Asn Pro Thr Ala Gln Pro Gln Ser 375 Amp Fro Thr Ala Gln Pro Gln Leu Asn Leu Met Ala Gln Pro Gln Ser Applier Val Ala Gln Pro Gln Ala Asp Thr Asn Val Gln Thr Pro Ala 410 Pro Ala Ala .:10. 95 12111 49 11111 DNA 213 Artificial Sequence :1201 misc feature Novel Sequence 012.231 -01000 95 thowardt atggaateat ettteteatt tggagtgate ettgetgte 49 11.100 96 1111 49 1.12 DNA #130 Artificial Sequence 0.201 <L21 rdsc_feature
<L23 Novel Sequence</pre> <400 > 96 4.9 ttcactcgag ttagccatca aactctgagc tggagatagt gacgatgtg Page 70

-(210)- -(211)- -(212)- -(213)-	97 22 DNA Artificial Sequence	
	misc_feature Novel Sequence	
-1001 gnt.caa	9/ coma oticatotatg co	2
- 210:- - 211:- - 213:- - 213:-	The state of the s	
<pre>// 2019/6- // 201</pre>	misc_feature Novel Sequence	
+400 + asact t	08 dict goodtaceg to	2
100 112 13		
::(0)- - 2211- - 223	rusc_feature Novel Sequence	
< 4(m); шаацса	99 goad coogaatadd	2
· 2100 · 2100 · 2100 · 2130 · 2130	41	
· 274) · · · 274) · · · · 223 · ·	nusc_feature Novel Sequence	
-400: :atgat	InO chae etgagegtea e	2
. 2105	3 (-1	

WO 01/36473

(211) (212) (213)		
	misc feature Novel Sequence	
∹400. ttoa∈a	1:1 gctt atggagtcgg ggctgctg	28
:000. :001: :013: :013:	107 30 DNA Artificial Sequence	
	misc_feature Novel Sequence	
::400° ttoact	102 cgag teagtetgea geoggttetg	30
0, 10 + 071 + 071, + 07		
0.1000 + -10.200 + -12.000 +	misc_feature N-wel Sequence	
:100 × goatic	ln3 tgde ogotatotgt goactotacg	30
0210 + 0211 + 0213 + 0213	30	
-:220 - -:331 - -:223 -	misc_feature Novel Sequence	
(40°) - Sątaya	1:14 gtgc acagatagog gocaggatgo	30
4::10 + (2:11 + (2:12 + (2:13 +	1.9	

4220>		
~321 •	misc_feature	
- 223	Novel Sequence	
· 400>	105	
aacccc	atca totacaego	19
	-	
1100>	106	
P. 11.15	18	
- 212>		
	Artificial Sequence	
	In Cliffic and Constitution of the Constitutio	
220.		
	misc feature	
	Novel Sequence	
	In ver Sequence	
- 40(C)	106	
		18
· unc.g	tuna geogetgg	10
- 711,00	107	
- 2111:-		
2119		
. 1.5,	Artificial Sequence	
2 Car		
	misc_feature	
. 22	Novel Sequence	
(; (; (;) ·		
grataa	gott ccatqtacaa egggtegtge tge	33
· _ 10. ·		
111		
12:		
- 213	Artificial Sequence	
- 2100		
_ 1 -	misc_feature	
- 223	Movel Sequence	
· 400	108	
uratto	tana teagtgeeae teaacaatgt ggg	33
(109	
- [1 1]	20	
	DNA	
- 1:13:-	Artificial Sequence	
	•	
2002		
	misc feature	
_		

-:2235	Novel Sequence	
-:400: gaagte	109 cage actgtttace	20
-1211: -212:-	1:0 20 LUA Artificial Sequence	
	misc_feature Novel Sequence	
-,⊈()(;} tqäääta	ll(acct gloogoagoo	20
- 2100 - 2115 - 2125 	35	
	nisc_feature Novel Sequence	
- 400: gut caai	111 gett atgacaggtg actteceaag tatge	35
111	112 34 MUA Artificial Sequence	
	Misc_feature Hovel Sequence	
.400. ⊋arokto	112 egag genaaeggea caaaacacaa ttee	34
10° 11 12 13		
	misc_feature Novel Sequence	

WO 01/36473	PCT/US00/31581
:400:- 113 :agcccaac atccaagte	19
02105 114 02117 19 0212 DNA 02138 Artificial Sequence	
02200- 02210 misc_feature 0223- Novel Sequence	
k4000 114 addocactta atcagoeto	19
+ 210% 115 + 211% 34 + 212% DNA + 213% Artificial Sequence	
- 220 - - 221 misc_feature - 223 Novel Sequence	
-4002 - 115 gatogaatho goaggagoaa tgaaaatoag gaac	34
+210+ 116 +211+ 39 +212+ DNA +213+ Artificial Sequence	
+220. +221 misc_feature 223. Novel Sequence	
:400°: 116 gatogaatto tiatatatgi toagaaaaca aattoatgg	39
+210 / 117 +111 / 10 +112 / DNA +213 / Artificial Sequence	
400% 117 magneceaa agedaaadad	20
*.10% 118 .11% 22 .212% DNA .213% Artificial Sequence	

PCT/US00/31581

(4-)0.5cogcage	ll8 gage aatgaaaate ag	22
1210 - 1211 - 1112 - 1113 -		
(40)0 - otgauad	119 gita tegetgace	19
:210. ::11. ::12. ::13.	DNA	
:1.20: :1.21- :1.25:	m_sc_feature Novel Sequence	
-:4u(: bgat fat	110 toda captitigado d	21
-1, 10 - -,:11; -1, 12 -,:13;		
dance poatace	721 catq aatgageeac tagae	25
	30 DHA	
	misc feature Novel Sequence	
:400 - gratuto	100 cqaq tcaagggttg tttgagtaac	30
1, 10 - 3, 11 - 4, 11 - 4, 15 -	20	
<.1.20 · <2.21 · <223 ·	misc feature Novel Sequence	

WO 01/36473	PCT/US00/3158

r14000 atgtat	123 ctet gteetettee	20
1210: 1211: 1212: 1213:	2.2	
3220. 121. 3223.	misc_feature Novel Sequence	
- 4000 udade g	124 atut toattgaatt to	22
- 210 211 212 213	CO.	
721. - 221. 223%	misc_feature Novel Sequence	
- 400c ⊶ctfca	175 aada actteatace ee	22
- 1:10:5 - 1:11:- - 1:11:- - 1:13:-	18	
	musc_feature Novel Sequence	
- 4000- acacac	126 aqda tagtagog	18
· 1301 · 211. · 212. · 23::	26	
	<pre>nisc_feature Novel Sequence</pre>	
4000	127	

cagagottga tgatgaggac		20
(110 ·		
	n.sc_feature Novel Sequence	
-(400 - coratac	1.0 gada gtagtagaag	20
<pre><210 + <211 + <212 + <213 + <226 + <221</pre>	1.9 q PET Synthetic substrate peptide m.sc feature	
	Novel Sequence	
	o Arg Thr Pro Gly Gly Arg Arg 5	
+210 + +211 + +212 + +213 +	5.1	
0220 · 121 · 4223 ·	misc feature Novel Sequence	
ावण्ण - प्राच्याख्य	130 ting actoactata gggagacogo gtgtotgota gactotattt co	52
*:210 / *:211 * *:212 * *:213 *		
	misc feature Novel Sequence	
<400→ tige pac	131 actg atgcaactcc	20

:211: :::12:	132 48 DNA Artificial Sequence	
	132 Lacg actoactata gggagacotg coacactgat gcaactec	48
**************************************	133 24 DNA Artificial Sequence	
(1) 100 (120 15 (100 30)	musc_feature Novel Sequence	
-400% quatato	133 etgc tagactotat itoo	24
-17101- -1.111- -1.122- -2131-	50	
- 4000- वृज्यtaat	104 tacq actcactata gggagacege acgceactet ttactatece	50
*2102 *2113 *2105 *2133	24	
	misc_feature Novel Sequence	
400° gcacuaa	$135_{ m dec}$ caattecata agec	24
110 ·	50	
	m.sc_feature Novel Sequence	
∴400 · qegtaat	136 zacq actcactata gggagacego acaaaacaca attccataag ee Page 79	52

<210: :211: :212: :13:	23	
	nisc feature Novel Sequence	
- 400) actaeg	107 ccac tottlactat ccc	23
1111	4.9	
	rasc_feature Novel Sequence	
edou grigtua	158 targ acticactata gggagacett atgageagea attratece	49
	159 10 DNA Artificial Sequence	
	misc_feature Novel Sequence	
-1400 + 	139 cauc aagaaatcag	20
1212	48	
	misc_feature Novel Sequence	
(400 · gogtia	140 staeg acteactata gggagaceca cacecaceaa gaaatcag	48
⟨210→	141	

·2115 · /121· · /135	21 DNA Artificial Sequence	
	misc_feature Movel Sequence	
	141 gray caattoatoo c	21
	142 49 ENA Artificial Sequence	
	misc_feature Novel Sequence	
- 400 gegtaat	142 acg actcactata gggagacccg attatccaca ctttgaccc	49
: 210 : : : : : : : : : : : : : : : : : : :		
	misc_feature Novel Sequence	
- 400 - etgaaag		19
· 2121	144 50 DNA Artificial Sequence	
	misc_feature Novel Sequence	
	144 Tacy actoactata gggagaccot gotgaaagtt gtogotgacc	50
::211 · ::212 ·	145 T1 DNA Artificial Sequence	

	misc_feature Novel Sequence	
-4002 udatta	145 tecă caetttgace c	21
.1100 .111. .112. .113.	50	
1220 1221 1223	misc feature Novel Sequence	
्द्रे()(- चुण्या तब	146 tacq actoactata gggagaccot gtaaaattoa cacaagouco	50
11.10 ⁴ · 12.11 · 12.11 · 13.1 · 13.1 · 13.1	1.4	
	misc_feature Novel Sequence	
:400 + adaada	14 ^{cc} caja gcaacotoo	19
-1210 - 1211 - -212 - -1213 -	1:8 48 DNA Artificial Sequence	
	nisc_feature Novel Sequence	
-:1))- dqoqta	148 atho gaeteactat agggagaeca gaagacagag caacetee	48
-:210 + -:211 = :212 -:213 +	2.2	
(220) + ≤221≥	misc_feature Page 82	

. 223>	Novel Sequence	
-400: ciqtaa	149 aatt cacacaagca cc	22
- 2105 - 2115 - 2125 - 2132	31	
	nisc_feature Novel Sequence	
-406- gostaga	150 atro totttgotgt atttoaccot o	31
- 210	21	
211 211 173	musc_feature Novel Sequence	
400 aratga	151 atto acaatgooag tgataaggaa g	31
- 2111- - 2121-		
	misc_feature Novel Sequence	
- 4000 gatcaa	152 gott ggaatgatge cetttigeea e	31
× 1101- - 1111- - 12121- - 12131-	2.3	
-:2205 -:2215 -:2235	misc_feature Novel Sequence	

PCT/US00/31581

WO 01/36473

<4005 153 gateetegag cateatteaa agtaggtgg	29
(210 · 154 (211 · 42 (212 · DNA (213 · Artificial Sequence	
<pre>ALLO + ALLO = misc_feature ALLO = Novel Sequence</pre>	
-400. 154 43 ontogeg statgaasté aattesaaaa ataatttasa so	42
0.100 155 -0.110 49 -0.110 FNA -0.130 Artificial Sequence	
Surfice FULL: masc_feature FULL: Masc_feature FULL: Masc_feature	
$^{\circ}4(.0) \leq 185$ $_{\odot}$ and the action of the state of	49
::00	
<pre>clift risc feature clift tovel Sequence</pre>	
- 4000- 156 - 4001 Agaaaa atgcataaac cattggatta aatgtagagt tcaagtagc	49
#210 - 157 5/11 - 35 #/12 - DNA #21:3 - Artificial Sequence	
C.Co. SZT: misc_feature C223 - Novel Sequence	
-:400 · 157 gategaatte atggacacta coatggaage tgace	35

Page 84

12125	158 31 DNA Artificial Sequence	
::[20: ::[21: ::[23:	misc_feature Novel Sequence	
	158 egag teaegtgggg cetgegeeeg g	31
-1210% -1211:- -212:-	$c_{1,2}$	
	misc_feature Novel Sequence	
- 4000- gogtaat	159 Laug actoactata gggagacogo gtgtotgota gactotattt oc	52
1101 111 112 113	20	
	misc_feature Novel Sequence	
- 400; « 1 genaca]w() actg atgcaactcc	20
1111	iul 4m DNA Artificial Sequence	
1200 2211 1230	musc feature Novel Sequence	
- 400g. acgtaat	161 acg acteactata gggagacetg ceacactgat gcaactee	48
2160	162 24	

:21.25 :213.	DNA Artificial Sequence	
	misc_feature Novel Sequence	
iajio. geatut	162 otgo tagactotat ttoo	24
1210.4 1211.4 1211.4 1213.5	5.0)	
1221 1221	misc_feature Novel Sequence	
∈4°0 - gogtaa	ୀରଃ tacq acteactata gggagaeege aegecaetet ttaetateee	50
-2102 -211 - -2.117 -2132	24	
	misc_feature Novel Sequence	
. ಕಲ್ಲೇ ಫಲಿಸಲಾಕ	164 aada caatteeata agee	24
100 -111 -112 -212 *217	50	
	nisc_feature Novel Sequence	
:400) gogtina	16% tadg actoactata gggagacogo acaaaacaca attocataag co	52
(210) (211) (212) (213)	23	

Page 86

<020: - 221: - 23:	<pre>misc_feature Novel Sequence</pre>	
-400. gatadge	166 scac tetttactat sec	23
- 210 211 212	4 ä	
121	misc_feature Novel Sequence	
- 400 g.gtaa	167 tacg actcactata gggagacett atgageagea atteateee	49
· 210 · · 2115 · · 2125 · 2132 ·		
+ 0000 + 021 + + 0234	misc_feature Novel Sequence	
- 4000 discard	168 cúcc aagaaatcag	20
11		
	n_sc_feature Novel Sequence	
- 400 ungtaa	169 taog actoactata gggagaccea cacceaceaa gaaateag	48
- 210:- - 211: - 212:- - 213:-	21	
· 220 · · 221 · · 223 ·	misc_feature Novel Sequence	

Page 87

-:40ō tta†qa	170 geag caatteatee c	21
0210 - 0211 - 0212 - 0213°	1/1 49 DHA Artificial Sequence	
+ 2, 61+ +(2,:11+ + 22+1	misc_feature Novel Sequence	
चेल्ह खब्दक्ष	101 tacq acteactata gggagaceeg attatecaca etttgacee	49
	19. 19 DNA Artificial Sequence	
	nisc feature Novel Sequence	
्कालए इ.च्छुट्टब्रह	172 gita tegetgace	19
11:10.14 (2:11.14 (2:11.14 (2:11.14)	50	
10.10	misc_feature Novel Sequence	
10406" 1041aa	1% tauq actoactata gggagaccot gotgaaagit gtogotgacc	50
(211)	174 21 DNA Artificial Sequence	
:220 :221 :223	misc_feature Novel Sequence	
<400 ·	174	

Page 88

cyatta	toca cactttgacc c	21
<pre>4310> 4211> 4212> 4213></pre>	50	
	m_sc_feature Novel Sequence	
-:400:- gogtaa	175 taog actoactata gggagaccot gtaaaattoa cacaagcacc	50
-0.10 - -0.11 - -0.12 2 - -0.13 -	19	
	misc_feature Novel Sequence	
- 40(I + adaāgā	176 caga gcaacctcc	19
*1210.+ *12117+ *12127+ *2133+	47	
1.110 · · · 1.11 · · · · · · · · · · · · · ·	misc_feature Novel Sequence	
- 400> gcgtaa	177 taog actoactata gggagaccag aagacagago aacetee	47
+210 + +111+ +212+ +113.+	211	
-11 0. -121 - -123 -	misc_feature Novel Sequence	
l4H0 - Htytua	178 aatt cacacaagca cc	22

WO 01/36473

4011 · 4212 ·		
120. 1221 123.	masc_feature Novel Sequence	
qüqe quatqga	179 afce tettigetgt atticaceet e	31
+1101+ +111.+ +2121+ +113 +	23	
	rdsc_feature Novel Sequence	
- ३००. नहत्व* वृक्क	180 atto acaatgocay tgataaggaa g	31
*210** *2111 *21** *21**	20	
	misc_feature Novel Sequence	
- 400, unageo	181 ocaa agodaaadad	20
+ 310% +,211. +(213)	44 44	
	misc_feature Novel Sequence	
	182 gago aatgaaaato ag	22
2.10 · 4211 · (212 ·)	2.0	

Page 90

WO 01/36473

· 213 ·	Arti	ficial	Sequ	ence				
		_featur Î Seque						
-400 - 1 ctgtatat	185 tot -	gteetet	tcc					20
-211 - 2 -212 - E	184 22 DNA Arti	ficial	Sequ	ence				
		_featur l Seque						
-400 - 1 gbackgat	184 toti	tcattga	att	tc				5.2
1112 I	185 1188 DNA H.Sa	piens						
::00	185			t	ogogaggta	cetagegage	cactageace	15 (a
						cctggcgagc		136
						tgqtactggc		
						ctgageteeg		180
						tgctggcggc		240
docttica:	ege	tgataga	gtgt	gatgcgcggg	eggacacegt	eggegeeegg	cgcatgccaa	300
greattg	gat	teetgga	acac	cttcctggcg	tecaacgegg	cgctgagegt	ggeggegetg	360
agegeag.	acc	agtggct	tggc	agtgggcttc	ccactgcgct	acgccggacg	cctgcgaccg	420
ogstatg	ccg	geetget	tgct	gggetgtgee	tggggacagt	agatggaatt	ctcaggcgct	480
geachtg	gct	gatagta	ggct	tggctacage	agegeetteg	egtectgtte	getgegeetg	540
engadag	age	ctgage	gtcc	gegettegea	geetteaceg	ccacgeteca	tgccgtgggc	600
ttogtgo	tga	cgctgg	eggt	gatatgaata	acctegetee	aggtgcaccg	ggtggcacgc	660
agacact	gcc	agegea'	tgga	caccgtcacc	atgaaggege	tegegetget	egeegaeetg	720
						ggcgccgcca		780
						ttgccccgta		840

Page 91

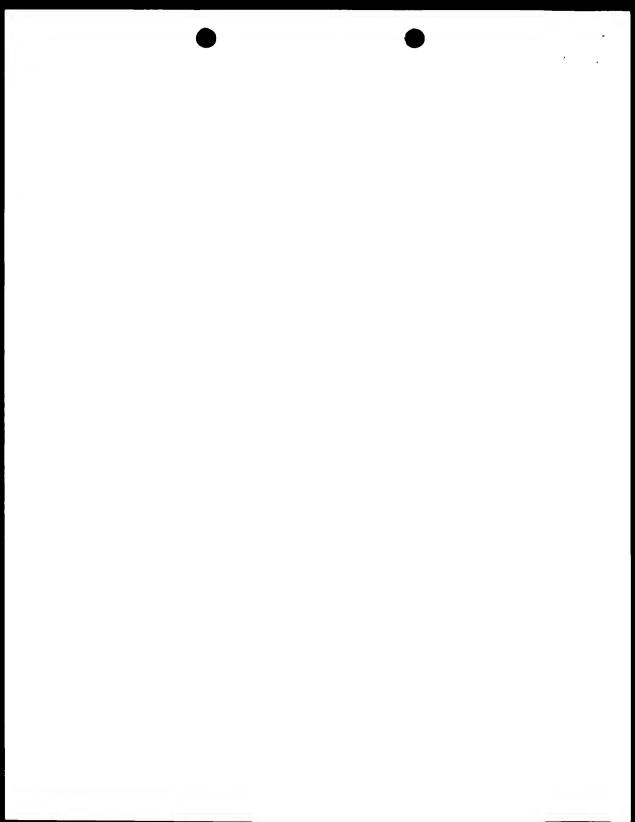
aggetggegg agetegtgee ettegteace gtgaacgeec agtggggeat ceteagewag
tgcctgacct acagcaagge ggtggeegae cegtteaegt actetetget cegeeggeeg
ttccgccang tootggccgg catggtgcac cggstgctga agaguaccec gegcccagca
tocacceatg acagetetet ggatgtggee ggeatggtge accagetget gaagagaace
ecquipecag egtecaceea caacggetet gtggacacag agaatgatte etqeetgeag
caganacact gagggeetgg cagggeteat egeocceaec tretaaga
<pre><210 - 186 <2211 - 363 <2212 - PET <2213 - H.Sapiens</pre>
-:400 · 186
Met Gly Fro Gly Glu Ala Leu Leu Ala Gly Leu Leu Val Met Val Leu 1 5 10
Ala Val Ala Leu Leu Ser Asn Ala Leu Val Leu Leu Cys Cys Ala Tyr $-25 - 30$
Ser Ala Glu Leu Arg Thr Arg Ala Ser Gly Val Leu Leu Val Asn Leu 35 40 45
Ser Leu Gly His Leu Leu Leu Ala Ala Leu Asp Met Pro Phe Thr Leu 50 60
Lou Gly Val Met Arg Gly Arg Thr Pro Ser Ala Pro Gly Ala Cys Gln 65 70 75 80
$7a1$ The Gly Phe Leu Asp Thr Phe Leu Ala Ser Asn Ala Ala Leu Ser $-\frac{35}{90}$
Val Ala Ala Leu Ser Ala Asp Gln Trp Leu Ala Val Gly Phe Pro Leu 100 105 110
Arg Pyr Ala Gly Arg Leu Arg Pro Arg Tyr Ala Gly Leu Leu Gly 115 120 125
Cys Ala Trp Gly Gln Ser Leu Ala Pne Ser Gly Ala Ala Leu Gly Cys 130 135 140
Ser Trp Leu Gly Tyr Ser Ser Ala Phe Ala Ser Cys Ser Leu Arg Leu 145 150 150 160
Pro Pro Glu Pro Glu Arg Pro Arg Phe Ala Ala Phe Thr Ala Thr Leu 165 170 175
His Ala Val Gly Phe Val Leu Pro Leu Ala Val Leu Cys Leu Thr Ser 180 185 190

Leu Gln Val His Arg Val Ala Arg Arg His Cys Gln Arg Met Asp Thr

Page 92

Val	Thr 2:0	Met	Lys	Ala	Leu	Ala 215	Leu	Leu	Ala	Asp	Leu 220	His	Pro	Ser	Val	
Arg 225	Gln	Arg	Cys	Leu	Ile 230	Gln	Gln	Lys	Arg	Arg 235	Arg	His	Arg	Ala	Thr 240	
Arg	Lys	Ile	Gly	Ile 245	Ala	Ile	Ala	Thr	Phe 250	Leu	Ile	Cys	Phe	Ala 255	Pro	
lyr	Val	Met	Th.r 260	Arg	Leu	Ala	Glu	Leu 265	Val	Pro	Phe	Val	Thr 270	Val	Asn	
Ala	Gin	Trp 275	Gly	Ile	Leu	Ser	Lys 280	Суѕ	Leu	Thr	Tyr	Ser 285	Lys	Ala	Val	
Ala	Asp 130	Pro	Phe	Thr	Tyr	Ser 295	Leu	Leu	Arg	Arg	Pro 300	Phe	Arg	Gln	Val	
Leu 205	Ala	Gly	Met	Val	His 310	Arg	Leu	Leu	Lys	Arg 315	Thr	Pro	Arg	Pro	Ala 320	
Ser	Thr	His	Asp	Ser 325	Ser	Leu	Asp	Val	Ala 330	Gly	Met	Val	His	Gln 335	Leu	
L∈u	Lys	Arg	Thr 340	Pro	Arg	Pro	Ala	Ser 345	Thr	His	Asn	Gly	Ser 350	Val	Азр	
Thr	Glu	Asn 355	Asp	Ser	Cys	Leu	Gln 360	Gln	Thr	His						
- 21 - 21 - 21 - 21	1 2:	187 29 INA Arti	fici	al S	eque:	nce										
	1:	misc Nove		ture quen	ce											
	0⊹ taag		gcca	tggg	cc c	cggc	gagg									29
· 21 · 21 · 21 · 21	2:	188 18 LNA Arti	fici	al S	eque	nce										
422 422 422	1:			ture												
	00> attot		cctc	agtç	ıtg t	.ctga	tgc									28
<2.	10>	189														

- 212>	20 DNA Artificial Sequence	
	misc_feature Novel Sequence	
(40)(i) rigot got	189 titg tigogootac	20
+02100+ :02115+ +02125+ +02133+	18	
	misc_feature Novel Sequence	
:400> ttggac	190 gcca qgaaggtg	18



SECUENCE LISTING

<110> Pharmacia & Upjohn Company Vogeli, Gabriel Hoff, Rita Sejlikz, Torsten Ling, Peter Slightom, Jerry Schellin, Kathleon Bannigan, Chris Buff, Valerie Raytes, Paul Wood, Linda Parodi, Luis Hiebach, Ermald <1200 Novel & Protein Coupled Receptors <130> 043F1PHRM296 <150> 60/165,830 1999-01-16 <1512 <1505 60/198,568 <194> 2000-04-20 <150> 60/166,071 <151> 1999-11-17 <150> 60/166,676 <151> 1989-13-19 <150> 60/173,396 <151> 1999-12-28 <150> 60/184,129 <151> 2000-02-22 <150% 60/185,421 <151> 2000-02-28 <150> 60/085,554 <151> 2000-82-26 <150> 60/186,590 <151> 200**0**-03-02 <150> 60/186,811 <151> 2000-03-03 <150> 60/188,114 <151> 2000-03-09 <150> 60/190.310 <151> 2000-03-17 <150:- 60/190,800 <150> 2000-03-21

Page 1

```
<150>
       60/201,190
<151>
       2000-05-02
<1.5.0>
       60/203.113
<151>
       2000-05-08
<1.50>
       60/207,094
<151> 2000-05-25
<160%
       190
<1700 PatentIn version 3.0</p>
<210:-
       1182
:211:
(212>
      [NA
:213>
      H.Sapiens
4400>
othtogogit gagagatget aggasagaga teaattgeet gaegsaegta eteteetees
                                                                       60
cotagetest getgatetag tiggigetes agagtgygga ggøgexeyge achtigaeas
                                                                      120
                                                                      160
ttototacno ttacoquelt ageceleass etetesgets gasatastes egatetesa
                                                                      240
rogacttter studenstat studecaca attoctqqcc qagaqaaaqa ggagqaatqa
egygageane itetteacte etagggeest gtggtagage tgeagteges ectectiotg
                                                                       300
                                                                       360
ceaataqqee taqatqaqtq qqttqaqqaq qqaqttqooc qoqqoqoqqo qoqacoqqta
                                                                       420
controler actarguage antiqueste obquesque acetquesse teccaquas
аждушадууд уборыунака узургаздек состатуада асауасгону каруланас
                                                                       4800
                                                                      540
tttungalen etaggagtes gtggggateg ataacetesa gesatggete stgsatgtte
                                                                       600
esuctifics attigotogic totopatoga oppaatotic ageatotoco agtaçaagaa
pacakagang agcatogoto opaspaagoo wacqoaqqaq aqqqtoaqow oqaactgacq
                                                                       660
                                                                      720
gigazatada gegazgonge igeaciquee ittüqiaqqon qiciquiqqq acabeqqqqui
toeqaqtiqeq aqqaaqoouu biqaqqtaaqa cactaaacao aqoccqquaa tqcaqqqcoo
                                                                      フリハ
ggopaggasa cosolastge Lottnasgta goggasgggo tgottgatgg casggtacct
                                                                      840
wtopassyty atcagcatga ocytyaggac agagycagot goggaggaag tgacaaatyc
                                                                       906
estecchago otopacagos tottototot ogocopagaa gogotochoa octoptotot
                                                                      960
gagtaggoca gagatggoca caccaatoao qqtqtcaqoc acoqueeqat toaaqqtqaa
                                                                      1020
quagagnety achocatest telligityyst esacagesgo acagedanay ecactagist
                                                                      1080
nttaqtomou stqslijaggy aggeesgyae egesaggate acteessaty agasagatga
                                                                      1140
```

Page 2

1182

ttocatqtok ogaagtogoa qgabttbabt tabbaqqqba tq

<210> 2 <211> 335 <212> PRT <213> H.Sapis	as			
<400> 2				
Met Glu Ser Se 1	r Phe Ser Ph S	he Gly Val	lle Leu Ala 10	Val Leu Ala Ser 15
Len Ile Lie Al 20	s Thr Asn Th	hr Leu Val 25	Ala Val Ala	Val Leu Leu Leu 30
lle Hi≥ Lye Ae 35	n Asp Gly Va	al Ser Leu 40	Cys Phe Thr	Lau Asn Leu Ala 45
Val Ala Asp Th 50	r Leu Ila Gl 55	ly Val Ale S	lle Ser Gly 60	Leu Leu Thr Asp
Gin Wee Ser Se 65	r Pro Ser Ar 70	rg Pro Thr	Glo Lye Thr 75	Leu Cys Ser Leu 80
Arg Met Ala Ph	e Val Thr Se 85	er Ser Ala	Ala Ala Ser 90	Val Lou Thr Val 95
Met Leu Ile Th		rg Tyr Neu 105	Als The Lys	Gin Pro Phe Ang 110
Tyr Leu Lys II 115	e Met Ser Gl	ly Phe Val 120	Ala Gly Ala	Cys Ile Ala Gly 125
Leo Trp Leo Va 130	l Ser Tyr Lo 13	ou 110 Gly 35	Pho Lou Pro 140	Less Gly IVe Pro-
Met Phe Gln Gl 145	n Thr Ale Ty 150	yr Lys Giy	Glm Cys Ser 155	Phe Phe Ala Val 160
Phe His Pro Hi	s Phe Val Le 165	en Thr Lev	Ser Cys Val 170	Gly Phe Phe Pro 175
Ala Mot Lou Le 18		he Pbe Tyr 185	Cys Asp Met	Leu Lya Ile Ala 190
Ser Met His Se 195	r Glm Glm II	le Arg Lys 200	Met Glu His	Ala Gly Ala Met 205
Ala Gly Gly Ty 210	r Arg Sor Fi 20	ro Arg The 15	Pro Ser Asp 220	Phe Lya Ala Leu
Arg Thm Val Se 225	r Val Leu II 230	le Gly Ser	Phe Ala Leu 235	Ser Trp Thr Pro 240
Phe Lew Ile Th	r Gly Tle V a 245	al Glm Val	Ala Cys Gln 250	Glu Cys Bis Leo 255
Tyr Leu Val Le 26	u Glo Arg Ty O	yr Leu Trp 265	teo Leo Gly	Val Gly Aen Ser 270

Leo Lao Ash Pro Leo Ile Tyr Ala Ty: 275 280	r Trp Glm Lya Glu Val Arg Leu 285
Glm. Let Tyr His Met Ala Lou Gly Val 290 295	l Lys Lys Val Lou Thr Ser Pho 300
Leo Leo Phe Leo Ser Ala Acg Asn Dy. 305 310	s Gly Pro Glu Arg Pro Arg Glu 315 - 320
Ser Ser Cys His lie Val Thr lie Set 325	r Ser Ser Glu Phe Asp Gly 330 335
<210> 4 <211> 697 <212> DNA <213> H.Sapiena	
<400 > 3	etocurtza cantutrino atascacca 60
cagonogaço goetteabyo tqacqqtqLc c	begindenda codebecte debecent
gtocaucton agonaegtom ggesgagese c	approduction agreed and a careful and a
dańcd.ddod breweddeta cdwydeddd w	
რეოდყორები aaggegetge tgtagecaag c	
gyccaycyae tgtebccagg cacagoccag c	ayoaygoog goerageyog gerge—ggra
teeggegtag egeagtggga ageccaelge s	
gotoapogoo gopubngang nnaggaaggt g	
gagoacogse gatatecace egegeatear a	
egocageage aggtggooda gaqxxxqqatt o	
gayotoaqog otgtaqqoqo aacaaaqoaq o	
cagtaccate accaggagas copecageas c	goekeyeeg gagecen tuu agelage 657
<210> 4 <211> 217 <212> PRT <213> H.Sapiens	
<400> 4	
Ser Ala Met Gly Pro Gly Glu Ala Le 1 5	m Leu Ala Gly Leu Leu Val Met 10
Val Leu Ala Val Ala Leu Leu Ser As 20 25	m Ala Leo Val Leo Leo Cye Cye 30
Ale Tyn Ser Ale Glu Leu Arg Thr Au 35 40	g Aka Ser Ciy Val New Lew Val 45
Aen Leu Ser Leo Gly His Loo Leo Le 30 55	en Ala Nia Leo Asp Met Pro Phe 60
	Page 4

Thr Leu Leu Gly Val Mot Ang Gly Arg Thr Pro Ser Ala Pro Gly Ala 65 70 80	
Cys Glm $V_{\rm ell}$ lie Gly Phe Leo Asp Thr Phe Leo Ala Ser Asn Ala Ala 85 -90 95	
Leu Ser Val Ala Ala Leu Ser Ala Asp Glm Trp Leo Ala Val Gly Pho 100 105 110	
Pro Leu Arg Tyr Ala Gly Arg Leu Arg Pro Acg Tyr Ala Gly Leu Leu 115 120 125	
Leu Cly Cys Als Trp Gly Gin Ser Leu Als Phe Ser Gly Als Als Leu 130 135 140	
Gly Cyp Sox Trp Les Gly Tyr Ser Ser Ala Phe Ala Ser Cys Ser Leu 145 150 155 160	
Arg Leu Pro Pro Glu Pro Glu Arg Pro Arg Phe Ala Ala Pho Thr Ala 165 170 175	
Thr Leu Bis Ala Val Gly Phe Val Leu Pro Leu Ala Val Lago Cys Leu 180 185 190	
Thr Ser Lou Gln Val His Arg Val Ala Arg Arg His Cys Gln Arg Met 195 200 205	
Asp Thr Val Thr Met Lys Ala Leu Ala 210 215	
<210> 5 <211> 222 <212> DNA <213> B.Sapiens	
:400> 5 tgtgcaggtg lgateteest teetttgtae atoopteaca egetgttega atgggatitt	6D
qqaaaqqaaa totgtqtatt ttqqotoact actgactate tqttotqtac accatotqta	120
tstascatty tecteateay etatgateya tacctyteay teleaaalige tytaaglega	180
acacattast ttmtcoccct baqsayatta tgtsaatgta ts	222
<400> 6	
Cya Ala Gly Val Ilo Sor Ile Pro Leu Tyr Ile Pro His Thr Leu Phe 1 5 10	

Page 5

Tyr Leu Ces Thr Ala Ser Vol Tyr Ash Ile Val Leu Ile Ser Tyr 35 40 45

Asp Ang Tyr led Ser Val Ser Asm Ala Val Ser Arg Thr His Phe Ils 50 55 60

Pro Leo Arg Arg Leo Cys Lys Cys Ile 65 70

<210> 7 <211> 507

<21.5 DMA

<2135 H.Sapiens

<400> 7 gangtogaag caggigaiga igeonagggo gigeacoggg iaggigagai oggigogogo 60 cagoggigas agggoggtos ggagsagsag scaggtocot gcasacgogg coaccquyta 120 acqaeqqeqq egccaqeget tqqaqetqag eggqtacagg atocccaqqa agegetecae 180240gotpatacay gtoathygiga gwaihothyga atacaigitt gogiaaaagg ccaeggicae 300 cacquingosa agosqosoco ogastacoca gtgqtqqcgq ttgcaatgqt agtagatttg qaaaggusar abgulqguca geateaggtu ogbgangele aggttgatea tgaagatgae 360 севордары, отдресска трееструсы сырсысскый междинамую прітрессорд 420 480 gatgotgaes geogocacca gegagtacae caegggeagg gecacegega tegeogggbb 507 cogeageate tgeagegteg egitigte

<\$10> B

<111> 169

<212> PST

<213> H.Sapiens

<400> 8

Asp Asm Ata Thr Leo Glo Met Ven Arg Asm Pro Ala Ile Ala Val Ala 1 5 10

Leo Pro Val Val Tyr Ser Leo Val Ala Ala Val Ser Ile Pro Gly $\hbar \sin 20$ 25 30

Leu Phe Ser Leu Trp Val Leu Cys Arg Arg Met Gly Pro Arg Ser Pro 35 40 45

Ser Val lie Phe Met lie Ash Leu Ser Val Thr Asp Leo Met Leo Ala 50 55 60

Ser Val Leu Pro Phe Gin lie Tyr Tyr His Cys Aso Arg His His Trp 89 78 80

Val Phe Gly Val Leu Cys Aso Leu Val Val Thr Val Ala Phe Tyr Ala 85 90 95

Ash Met Tyr Ser Ser Ile Leu Thr Met Thr Cyo Ile Ser Val Glo Arg 100 105 110 Phe Leu Gly Ile Leu Tyr Pro Leu Ser Ser Lys Arg Trp Arg Arg 115 120 125

Mrg Tyr Ala Val Ala Ala Cys Ala Gly Thr Trp Leu Leu Leu Leu Thr 130 140

Als Leu Ser Pro Leu $\Lambda 1\sigma$ Arg Thr Asp Leu Thr Tyr Pro Val His Ala 145 155 160

led Gly Ile The Thr Cys Fhe Asp Val

cgcacyctog cgatygcago cgceecctee

<210> 9 <211> 270 <212> DNA <213> H.Sapiens

<400> 9
cocatgites igeteciags casceleacy instrugeats byethyseasy egospecias 60
geogrosasea isetactific gaggeogete accetyaaas tylonoonge geteigette 120
geonogagaag gaggestett eqisgeacic actgestees igetgagent oologgeote 180
geogrosasea, qeageotean categospen agggggooon egocosteic castegospe 240

270

<210> 10 <211> 90 <212> PRT <213> H.Sapiens

<400> 10

Pro Met Phe Leo Leo Gly Sor Leo Thr Leo Ser Asp leo Leo Ala 1 5 10 15

Gly Ala Ala Tyr Ala Ala Asn Tle Leo Leo Ser Gly Pro Loo Thr Leo 20 25 30

Lys Lou Ser Pro Alo Leu Trp Phe Alo Arg Glu Gly Gly Val Phe Val 35 40 45

Ala Leo 3hr Ala Ser Val Leo Ser Leo Gly Ile Ala Leo Glu Arg 50 60

Ser Leu Thr Met Ala Arg Arg Gly Pro Ala Pro Val Ser Ser Arg Gly 65 70 75 80

Arg Thr Leu Ala Met Ala Ala Ala Ala Trp 85

<210> 11 <211> 600

<212>	DNA
<213>	H. Sapiene

<400% 11

<400> 11						
	tggcotttgt	getgggegea	obaggdaat,g	gggtogcoct	gtgtggtttc	60
tgottocada	tgaagacotg	gaageccage	actgtttacc	Litteastit	ggongtgget	120
gatttestes	ttatgatotg	cotgootttt	oggacagact	attecctcag	adgtagação	180
üeggatttta	gggscattcc	otgrogagtm	gggetettea	cgttggccat	gaacagggcc	240
gggagcatcg	tgttccttac	ggtggtggct	geggaeaggt	attiteanøgt	ggtocacocc	300
caccacgegg	tgsacactat	ctocaccogg	gtagaagaata	geategicing	cacontatas	360
quoutegton	tcotgggaac	agtgtatett	ttgotggaga	accatetetg	cgtqcasqag	420
acggoogtet	cetutqagaq	ottoateatg	gaytoggoda	atggctggca	tgacatcatg	480
ttocagetge	archictitat	quocatagga	atoatettat	tttgetoett	caagattgtt	540
tggagentga	decdéadace	geagetgeer	ადასახვიზი	gņotģaagaa	ggcqacccgg	600
ttoatcatgy	tçgtçgcaat	tgtgttcatc	acatgctacc	tgoocsqogt	qtotqotaga	660
etetatitee	totagaoggt	goddtogagt	geetgegate	cototytens	Lgggggggttg	720
oacataacco	teagotteac	ctacatgaac	agcatgotgg	atcccctggt	gtattetitt	780
heaagemeet	ectificacea	abtotacaac	aaget caeaa	totgcagtot	gaaacccaag	НАП
падорацуяс.	ectoeaaeec	асаваддоор	gaaqagatge	caatttcg		888

<400> 12

Leo Leo Ile Val Ala Fhe Val Leo Gly Ata Loo Gly Asm Gly Val Ala

Let Cys Gly Phe Cys Phe Ris Met Lys Thr Trp Lys Pro Ser The Wal

Tyr Leu Phe Asn Lou Ala Val Ala Asp Phe Leu Leu Met Ile Cys Len

Pro Phe Arg Thr Asp Tyr Tyr Len Arg Arg Arg His Trp Ala Phe Gly $50\,$

Amp The Pro Cys Arg Val Gly Leu Phe Thr Leu Ala Met Asm Arg Ala

Cly Ser Ile Val Pho Leu Thr Val Val Ala Ala Asp Arg Tyr Phe Lys

<210> 12 <211> 296 <212> PRT

<213> H.Sapiena

WO 01/36473

### Fig. 11c Val Cys Thr Leu Trp Ala Leu Val Ito Leu Gly Thr Val 115	Val Val	Bis Pro 100		His	Ala	Val	Asn 105	Thr	Ile	Ser	Thr	Arg 110	Val	Ala	
Cys Glu Ser Phe lie met Glu Ser Ala Ash Gly Trp His Asp Ile Met 145 Phe Cln Leu Glu Phe Phe Met Pro Leu Gly Ile Ile Leu Phe Cys Ser 165 Phe Lys Ile Val Trp Ser Leu Arg Arg Arg Gln Gln Leu Als Ary Gln 185 Ala Arg Met Lys Lys Ala Thr Arg Phe Ile Met Val Val Ala Ile Val 195 Phe Lys Trp Trp Cys Tyr Leu Pro Ser Val Ser Ala Arg Leu Tyr Phe Leu 215 Trp Thr Val Pro Ser Ser Ala Cys Asp Pro Ser Val Mis Gly Ala Law 220 Trp Thr Val Pro Ser Ser Ala Cys Asp Pro Ser Val Mis Gly Ala Law 220 His Ile Thr Leu Ser Phe Thr Tyr Met Asa Ser bet Leu Asp Pro Leu 255 Val Tyr Tyr Phe Ser Ser Pro Ser Phe Pro Lys Gln Pro Gly His Ser Lys Thr Gln 285 Lys Ile Cys Ser Lou Lys Pro Lys Gln Pro Gly His Ser Lys Thr Gln 210	Ala Gly		Cys	Thr	Leu		Ala	Leu	Val.	110	Leu 125	@17	Thr	Va.1	
Phe Cln Leu Clu Phe Pho Met Pro Leu Gly IIe IIe Leu Phe Cys Ser 185		Leu Leu	Glu	Aen		Lėu	Сув	Val	Gln		Thr	Ala	Val	Ser	
The Lys lie was trop Ser Leu Arg Arg Arg Gln Gln Leu Ala Arg Gln 185 Ala Arg Met Lys Lys Ala Thr Arg Phe lie Met Val Val Ala lie Val 195 Pho lie Thr Cys Tyr Leu Pro Ser Val Sar Ale Arg Leu Tyr Phe Leu 215 The Thr Val Pro Ser Ser Ala Cys Asp Pro Ser Val His Gly Ala Luu 240 His lie Thr Leu Ser Phe Thr Tyr Met Asan Ser Met Leu Asp Pro Leu 255 Val Tyr Tyr Phe Ser Ser Pro Ser Phe Pro Lys Pro Lys Phe Tyr Aen Lys Leu 270 Lys lie Cys Ser Leu Lys Pro Lys Gln Pro Gly His Ser Lys Thr Gln 281 Arg Pro Glu Glu Met Pro lie Ser 295 4210 13 4210		Ser Phe	Ile		Glu	Ser	Ala	Asn		Trp	нів	Aap	Ile		
Ala Arg Met Lya Lya Ala Thi Arg Phe Ile Met Val Val Ala Ile Val 200 Phe Ile Met Val Val Ala Ile Val 200 Phe Ile Met Val Val Ala Ile Val 200 Phe Ile Thir Cys Tyr Lev Pro Ser Val Sar Ala Arg Leu Tyr Phe Leu 215 Phe Thir Val Pro Ser Ala Cys Asp Pro Ser Val Nis Gly Ala Lwo 240 Phe Ile Thir Leu Ser Ala Cys Asp Pro Ser Val Nis Gly Ala Lwo 240 Phe Ile Thir Leu Ser Phe Thir Tyr Met Asa Ser Det Leu Asp Pro Leu 255 Phe Ile 260 Phe Thir Tyr Met Asa Ser Det Leu Asp Pro Leu 255 Phe Ile 260 Phe Ile 260 Phe Tyr Ara Lya Leu 270 Phe Ile 270 Phe Pro Lys Gla Pro Gly His Ser Lys Thir Gla 271 Phe 275 Phe 295 Phe Ser Ser Pro 1 Pro Ser Phe Pro Lys Pro Ser Lys Thir Gla 280 Phe 1 Pro Ile Ser 295 Phe Pro Ile Pr	Phe Cln	Leu Glu		Pho	Me t	Pro	Leu		Ile	11e	Leu	Ph∈	Суа 175	Ser	
Pho lie Thr Cys Tyr Lev Pro Ser Val Sar Als Are Leu Tyr Phe Leu 215 Typ Thr Val Pro Ser Ser Als Cye Asp Pro Ser Val His Gly Als Lou 240 His lie Thr Leu Ser Phe Thr Tyr Met Ass Ser Met Leu Asp Pro Leu 255 Val Tyr Tyr Phe Ser Ser Pro Ser Pro Lys Phe Tyr Ash Lys Leu 275 Lys lie Cys Ser Leu Lys Pro Lys Gln Pro Gly His Ser Lys Thr Gln 290 <pre></pre>	Phe Lys		Trp	Ser	Leu	Arg		Arg	Gln	G1n	Leu		Vid	Gln	
The Thr Val Pro Ser Ser Als Cys Asp Pro Ser Val Ris GLy Als Loc 240 His The Thr Leu Ser Phe Thr Tyr Met Asm Ser Met Leu Asp Pro Leu 255 Val Tyr Tyr Phe Ser Ser Pro Ser Phe Pro Lys Phe Tyr Asm Lys Leu 255 Lys The Cys Ser Leu Lys Pro Lys Gln Pro Gly His Ser Lys Thr Gln 275 Arg Pro Glu Glu Met Pro The Ser 295 Arg Pro Glu Glu Met Pro The Ser 295 400 13 4212	Ala Arg		Lуз	Ala	ıdT	A rg 200	Phe	Ile	Met	Val	Val 205	Ala	ïle	Val	
His lie Thr Leu Ser Phe Thr Tyr Met Asn Ser Met Leu Asp Pro Leu 255 Val Tyr Tyr Phe Ser Ser Pro Ser Phe Pro Lys Phe Tyr Asn Lys Leu 255 Lys lie Cys Ser Leu Lys Pro Lys Gin Pro Gly His Ser Lys Thr Gin 275 Arg Pro Glu Giu Met Pro lie Ser 295 <pre> </pre> <pre> </pre> <pre> </pre> <pre> </pre> <pre> </pre> <pre> <pre> <pre> </pre> <pre> <pre> </pre> <pre> <pre> <pre> </pre> <pre> <pre> <pre> <pre> <pre> </pre> <pre> <pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre>		Thr Cys	Tyr	Leu		Ser	Val	Ser	Ale		Lea	Туг	Phe	Leo	
Val Tyr Tyr Phe Ser Ser Pro Ser Phe Pro Lys Phe Tyr Asn Lya Leu 270 Lys Ile Cye Ser Leu Lys Pro Lys Gin Pro Gly His Ser Lys Thr Gin 275 Arg Pro Glu Glu Met Pro Ile Ser 295 <pre> </pre> <pre> </pre> <pre> </pre> <pre> </pre> <pre> <pre> <pre> <pre></pre></pre></pre></pre>		V al Pro	Ser		Ala	Сує	Дер	6, o		Val	His	GLy	А). а	Two 240	
Lys lie Cys Ser Leu Lys Pro Lys Gin Pro Gly His Ser Lys Thr Gin 275 Arg Pro Glu Glu Met Pro lie Ser 295 <pre> </pre> <pre> <pre></pre></pre>	His lle	Thr Leu		Phe	Thr	Тух	Met		Ser	het	Геп	Aep.		Leu	
Arg Pro Glu Glu Met Pro Ile Ser 295 <pre> <pre></pre></pre>	Val Tyr		Ser	Ser	Prc	Ser		Pro	Lys	₽he	Tyr		гћа	ren	
<pre></pre>	Lys Ile	Cys Sor 275	Lou	Lys	Pro	Lys 200	Gln	Pro	ĠΙγ	Нав	Ser 205	Lys	Thr	Gln	
<pre> <211> 510 <212> DNA <2130 H.Sapiens c400> 13 tgggagetytg organization offgatyging objectively gradient that tiggets figuragetyte toatcatose offactoria gatyging gradient gatyging light gatyging gradient gatyging light gatyging gradient gatyging light gatyging gatygi</pre>		Glo Glo	Met	Pro		Ser									
transported coarcacta totagetano organizate connected that process 60 transported toatcaton oracleonia gargaenage agreening against 120 transported transported geteratate anomalization against transported generated 300 of the generated	<211> 3 <212> 0	510 NA	5												
tgcaagetgg tgcactteet gttetatate wacetttaeg geageateet getgetgace 180 Lgcatetelg tgcancagt: entaggtgtg tgccancoat tgtgtteget gecetacegg 240 accegoagge atgcetgget gggcaccage accacctggg contggtggt cotcoagetg 300 etgcccacse tggcettete ceacseggae tacateaatg gccagstgat etggtatgae 360 atgaccagee aagsgaattt tgateggett tttgcctaeg gcatagttet gacattgtet 420			cepot	ia t,c	:Ççgt	្ឋប្បទ	o nite	gatgi	gragg	ं ंव्य	acet.	got t	ttato	ıtgeta	60
Liquetetetiq tigaannagti ontaggtigt liqueannoon tigligticiget geociacogg 240 annogragge stycetiget gagescage accaectigg contiggtigt collectagetig 300 etigennacse tiggeritete ceansaggse tacateastig geragstigat etiggtaligae 360 stigaecagoe aagsgastit tigatoggott tittgeetaog gestagttet gaestigtet 420	ttgesett	ce teat	catca	ac ot	iacto	cacta	a gat	gada	aggt	ggaa	octti	egg (ggage	ntgeto	120
ennegragge stycetyget gggesecage accaectygg contygtggt coloragety 300 etgenhacse tygentete ceaeseggs taeateasty genagstgat etggtalgae 360 stysecage asysgastt tygetegett titgeetaeg genagstet gaesitytet 420	tgcaaget	igg tgca	ottoo	at gt	tota	atato	s eat	octti	tacg	geag	jeati	oot (gotga	etgacc	180
etgennacse tggenttete ceacseggse tacatematy genagstgat etggtalgan 360 stgannagen aagsgasttt tgateggett tttgentacy geatagttet gansttgtet 420	Ugeatete	eta tada	nnagt	it or	tagą	gt gto	g lgi	របស់ការ	() ((4)) ()	t,qL(ybbo	ցուն գ	good	acogg	240
atgaccagec aagagaattt tgateggett tttgeetaeg geatagttet gaeskigtet 420	accognag	gge atge	etgga	et ge	gges	ecag	2 800	eacc1	tggg	cent	tggti	ggt. i	ooke	ragety	300
	etgennac	ese tigge	ettet	0 00	racad	ogga:	o tad	cate	aatg	gee	agati	gat :	etggl	talgac	360
	atgaccaç	gcc aags	gaatt	it to	gate	ggat1	t tti				tagt	tat 1	gaca:	etgtet	420

480 510

ggotttettt ecotectigg teattitiggt gigotatica organizates ggageorgat													
casgocagas gagaacotoa tgaggacags													
<210> 14 <211> 170 <212> PRT <213> H.Sapions													
<400> 14													
Trp Ser Cys Ala Thr Thr Tyr Leu Val Asn Leu Met Val Ala Asp Leu 1 5 10 15													
Leu Tyr Val Leu Pro Phe Leu IIn The Thr Tyr Ser Leu Asp Asp 20 25 30													
Ang Trp Pro Phe Gly Glu Len Len Cys Lye Leu Val His Pho Len Phe 35 40 45													
Tyr lle Aan Leu Tyr Gly Ser Ile Leu Leu Leu Thr Cya Ile Ser Val 50 55 60													
His Gln Phe Lea Gly Val Cys His Pro Lea Cys Ser Lea Pro Tyr Arg 65 70 75 80													
The Arg Arg His Als Trp Leu Gly Thr Ser Thr Thr Trp Ala Leu Val 85 90 95													
Val Leu Gln Leu Leu Pro Thr Leu Ala Phe Ser Bis Thr Aap Tyr I)e 100 105 110													
Aen Gly Glm Met Ile Trp Tyr Asp Met Thr Ser Glm Glu Asn che Asp 115 120 125													
Arg Leu Phe Ala Tyr Gly Ilo Val Leu Thr Leu Ser Gly Phe Leu Ser 130 135 140													
Let Let Gly Nis Phe Gly Val Let Phe Thr Asp Gly Gin Glo Pro Asp 145 150 155 160													
Glm Ala Arg Gly Glu Pro His Glu Asp Arg 165 170													
<pre><2100 15 <211> 894 <2000 DMA <2130 W.Sapiens</pre>													
<pre><220> <221> misc_feature <222:</pre>													
:4003 15 composeggg cagcaegoog acagggcote teceteccat tetenogeag genegoacga													

PCT/US00/31581 WO 01/36473

ceacgetgee	tecamec eg t	oggosesota	aggeogeteg	оадосскота	acagoagoco	120
capcayetgg	oteatettea	ggabatgaaa	otthgogogg	ggoatogogo	tgggegeacg	160
gmotopacot	gggatagaag	accaggoogo	tgcacesquet	ggggeettea	gecggtgeeg	240
ncaccagaeq	qaqaqtaqqt	gg¢qacasq¢	gacecocate	atottaacag	dododecdas	300
daccdadwad	gcctcatega	acgegtacae	ctgcacgtgc	овредствея	gg#gogogaa	360
gatobagtgg	cagogaogos	tepeoggoea	ggstegggeg	gagagtggcg	ageatograto	420
cagagaegtit	аллаялаала	תתמתתמותממ	חתמממחתמממ	nagtactage	дсассасава	480
офестановае	ogenceages	qbaytqcceq	съдесивосс	အခဲ့ခဲ့ခဲ့ခဲ့ခဲ့ခဲ့ခဲ့ခဲ့ခဲ့ခဲ့ခဲ့ခဲ့ခဲ့ခ	gggeacgege	540
ggqdagdeigd	nggnoglang	ពួនគម្ពង៤ព្រះគ¢	ព ្ ក់កុំង្គប់អ្នកកក្នុង	equitegangg	egatgagsae	600
cacqaqqtqq	geegaggege	ecogocogga	tgeetgesge	agetgeeega	адордоводо	660
caggtoccoo	gtiggaegage	ggggatogaa	cagcagttcc	caggecaget	gtgaeagogo	7.20
agtmaccccag	cachegtaca	ggt eegeeag	gçocagotgo	ಪ ಾರವಿದ್ದರಿಕಾಗ್ಗಳಿತ	agtocatott	780
gegaeget.tin	ດການຄອກຄອກຄ	ממשתממת	эванликин	ងព្ ថ្មជាពុទ្ធព្រង(អ	guastytygt	840
gttgeetgee	adegecacea	ccaggatgac	ccccaggaac	accaggcgga	ogeş	H94

<4000> 16

Arg Val Arg Deo Val The Leo Gly Val the Leo Val Val Ala Val Ala

Gly Asn Thr Thr Val Leu Cys Arg Leu Xaa Xaa Xaa Xaa Xaa Xaa Xaa

Maa Maa Lys Arg Arg Lys Met Asp Phe Leu Leu Val Glo Leu Ala

Leo Ala Asp Leo Tyr Ala Cys Gly Gly Thr Ala Leo Ser Glo Leo Ala Page 11

<210> 16 <211> 296 <212> PRT <213: H.Sapiens

<2200

<221> UNSURE <222> (26)..(35) <223> Xas is unknown

<220>

<221> UNSURE <222> (144)..(154) <223> Xaa is Unknown

	50					55					ħШ						
Тгр 68	G1o	Lev	Геп	Gly	Glu 70	Pro	Arg	Ala	Ala	Thr 75	Gly	Asp	Leu	Ala	Cys 80		
Àrg	Phe	Leu	Gln	Leu 85	Leu	Gln	Ala	Ser	Gly 9 0	Arg	Gly	Ala	Ser	Ala 95	H1s		
Leu	Val	Val	100 100	He	Alə	Leu	Glu	Arg 105	Arq	Arg	Ala	Val	Arg 110	Гел	Pro		
ніз	Glγ	Arg 115	Pro	Leu	Pro	Alz	Arg 120	Ala	Leu	Alm	A).a	Leu 125	Cly	1.cb	Len		
Leu	Ata 130	Leu	Leu	Leu	Ala	Arg 135	Gly	Ser	Gly	Fhe	V≳1 140	Val	ሽድዓ	Туτ	Xae		
Хаа 145	Maa	se%	Xaa	Xaa	Xaa 150	БеX	Хаа	Хъа	Жаа	Thr 155	Ser	Leu	Gln	Pro	160 GJA		
Al a	Pro	Testu	Sec	A18 165	Arq	ЛLа	Фтр	Pro	61 y 170	Met	Arg	Arg	Cys	His 175	Trp		
Ile	Phe	Ala	Leu 180	Leu	Gln	Arg	ттр	ніе 185	val	Gln	Val	ፒኍェ	Ala 190	Phie	тул.		
Glu	Ala	Val 195	Ala	G1 y	Phe	Val	Ala 200	Pro	Val	Lys	Ila	Met 205	Gly	Val	Ala		
Cys	G1 y 210	Ris	Iæu	T.Bu	Sex	Val 215	Tγp	Τπρ	yrd	His	Arg 220	lea	Lys	Ala	Pro		
225					Yrp 230					235					∠ 4 0		
				245	Ala				250					255			
Leu	Gly	Leu	\$60 Pao	Phe	Va)	Gly	Cys	G10 265	T _i ch (i	Pro	Phe	Ala	Asp 270	Arg	Leu		
GLu	Mla	A1a 275	Ттр	Ser	Sec	Sly	029 085	АІА	Gly	610	Ψt.fi	G10 285	СІУ	Glu	Ala		
Leu	Ser 290		Cys	Cys	Als	Trp 295	Trp										
<21 <21 <21 <23	1> 2>	17 801 DNA H.Sa	pien	*													
<40 tot	0.> aeat	17 ttt	tato	tgaa	ct t	tgag	actg	t ga	aass	agaa	ggg	stgo	tge	otoa	ggccac	€	60
															agtoat	13	įo
															catogg	3.5	iO

ofitgatgacc	ptaceggtgg	ocgsacctoq	бассяддаас	ocat tgggga	aytaytyyaa	240
outgatgoda	tgystgctgg	Lgttgggcag	ខ្លួនចំនុងជាមួយ	രാന്ദ്രമുമരളാ	cecagaegat	300
geogaqqato	ctgsqggccc	ggeçedgggt	getetgeagt	ttggcgcgga	əcqqqtqtag	360
gatggccacg	tagogotoca	oget gaeggt	ggtgatgetg	ទទិក់ទ ្រង់ពិទទិស	09 88ពុល ខ១៧០	420
ggtotoaaag	agggeegtet	tgaagtagca	decesedada.	cogasceaga	aaggytagbt	480
gegocacato	tcatagasot	ccangggdcat	tocaaggage	aggaccagga	ggt cagagec	540
cqccaggatç	nagaggtagt	agttegterg	egte tteata	gootggtgct	gcagaatcac	600
сидоренског	aggacatitgo	caatgaccoc	naccapaaaa	attqqcacat	асассасада	660
ចនព្វថ្មថ្មម្មវិទ្ធ	asymsytogo	targeogeég	tacquagagg	aaggooagat	actootoggt	720
gotgttdagg	tgtttetgga	atggatette	tagtttetge	tggtagstcc	aqgaaquatt	790
etgaagtiti	tocatcoctg	a				801

<210> 18 <211> 249

<212> PRT <213> H.Sapiens

₹400> 18

Ser Gly Met Glu Lys bea Gin Asn Ala Ser Trp Ile Tyr Gln Gln Lys 1 10 15

Leu Glu Asp Pco Phe Gin Lys Wis Leu Asn Ser Thr Glu Glu Tyr Leu 20 25 30

Ala Phe Leu Cys Gly Pro Arg Arg Ser His Phe Phe Leu Pro Val Ser 35 40 45

val val Tyr Val Pro Ile Phe Val Val Gly Val Ile Gly Aso Val Leu 50 55 60

Val Cya Leu Val Ilo Leu Gln Ris Gln Ala Mat Lys Thr Pro Aso Thr 65 75 80

Tyr Tyr Leg Phe Ser Len Ala Val Ser Asp Leo Leo Val Leo Leo Leo 99 99

Gly Met Pro Leu Glu Val Tyr Glu Met Trp Arg Asn Tyr Pro Phe Leu 100 105 110

Phe Gly Pro Val Gly Cys Tyr Pho Lys Thr Ala Leo Pho Glo Thr Val 115 120 125

Cya Phe Ala Ser Ille Dee See The The The Val Ser Val Glu Arg Tye 130 140

V_S: Ale The Lew His Pro Phe Arg Ale Lys Leu Glm Ser Thr Arg Arg 145 150 155 160

Arg Ala Leu Arg Tle Leu Gly Ile Val Trp Gly Phe Ser Val Leu Phe Ser Leu Pro Asn Thr Ser Ile Bis Gly Ile Lys Phe His Tyr Phe Pro Asn Gly Ser Leu Val Pro Gly Sot Ala Thr Cys Thr Val Ile Lys Pro Met Trp lie Tyr Asn Phe Ile Ile Glm Val Thr Ser Phe Leu Fhe Tyr 215 Len Leo Pro Met Thr Val Ile Ser Val Leo Tyr Tyr Leo Met Ala Lou Arg Val Ser The Ala Gly Val Ala Gly 245 <210> 19 222 <211> <2122 DNA <213> H.Sapiens <400> 19 atcasgatga tittigetat egigesaatt attggatiit eesseteest eigissieee attgtotatg catttatgaa tgaaaactto aaaaaasaatg ttttgtotgo agtttgttat 120 tgoatagtua atawaacott etetocages caeaggeatg gasuttengg sattacaatg 180 222 aljunggaaga sagnaaagti timnotcaga gagaatonag ig <210> 20 73 <211> <212> PRT <213> H.Sapiens <400> 20 lle Lys Met lle Phe Ala lle Val Gho lle lle Gly Phe Ser Asn Ser The Cys Ash Pro The Val Tyr Ala Phe Met Ash Glu Ash Phe Lys Lys Agn Val Leu Ser Ala Val Cys Tyr Cys Ile Val Asn Lys Thr Phe Ser Pro Ala Gln Arg His Gly Asn Sec Gly Ile The Met Met Arg Lys Lys Ala Lys Ehe Ser Leu Arg Glu Asn Pro

60

Page 14

<210>

<211> <212> DNA

21 447

PCT/U800/31581 WO 01/36473

<213> H.Sapiens

<400> 21 gecacagoat	gcagttttct	gtagaattcc	actttgtatt	tgcacttgss	Çaayatçaqq	60
tatotggtga	emaggateae	cacatagaat	aggaaccgtg	aggtacatgt	ggatgtgcag	1.20
catggcacto	acaaatttgc	agaagggcag	cocaaacato	caagtottot	tgatgaçgta	180
ggt.caaçoga	aatogcactg	bcagcagaaa	aacgetqtgg	ассассасся	agttæatgac	240
ogodat yglig	ghosotgaco	gggtgttcat	tticaccagg	aggaaaagaa	tegaaatgac	300
achdaerdagd	ocyccaataa	gcactatgaa	qtaqeggctq	atteagtqçç	qlgkcactat	360
aggat ogcaa	gaggaattcc	tggaggtatt	gtggecagge	atacttgyga	egboacotgg	420
aggagaaaaa	gcaccagagt	aactgac				447

<1100 Z2
<1115 149
<1125 PRT
<1135 R.Sapiena</pre>

<400> 22

Val Ser Tyr Ser Gly Ala Phe Ser Pro Pro Gly Asp Phe Pro Ser Met

Pro Cly His Asm Thr Ser Arg Asm Ser Ser Cys Asp Pro Ile Val Thr

Pro Ris Len Ile Ser Len Tyr Phe Ile Wal Len Ile Gly Gly Len Wal

Gly val Ile Ser Ile Leu Phe Leu Leu Val Lys Wet Asn Thr Arg Ser

Val Thr Thr Mot Alm Val Ille Asm Low Val Val Val His Ser Val Phe

Les Les Thr Val Pro Phe Arg Les Thr Tyr Les Ile Lys Lys Thr Trp

Met Phe Gly Len Pro Phe Cys Lys Phe Val Ser Ala Met Leu His Tle

His Met Tyr Leu Thr Val Pro Ile Leu Cys Gly Asp Pro Gly His Glm

lle Pro His Lou Leo Glo Val Glo Arg Glo Ser Gly Ile Leo Glo Lys

Thr Ala Cys Cys Gly 345

<210> 23 <211> 222

WO 01/36473	PCT/US00/3158.

<212> DNA <213> H.Sapiens	
<4005 23 antgaccaag gtcagggcat ogsotgaggo tagaaggoca caggaaatgo cagtcaaggt — 60	0
gttggageet gasatagead etaccacaan attgacongq ggaagnggggg cangacongco — 12	0
agogaacaog gtoaqoagos coagtocatt goagagcaog gagagcaaca ogatggocca 180	O
camagocarg opgatgocco agettteasa gaggtaetes ea 22.	2
<pre><2108 24 <211> 74 <212> PRT <213> H.Sapiens</pre>	
<400 = 24	
Cys Clu Tyr Len Pho Gis Son Trp Gly Lie Arg Leo Ala Vai Trp Ala 1 5 10 15	
Hit Val Leu Ser Val Leu Cys Ash Gly Leu Val Leu Leu Thr Val 20 25 30	
Fig. Als Gly Gly Pro Als Pro Leu Pro Pro Val Lys Phe Val Val Gly 35 40 45	
Ala Ile Ala Gly Ala Aon Thr Leu Thr Gly Ile Sor Cys Gly Leu Leu 50 55	
Ala Sor Wal Asp Ala Leu Thr Leu Wal Ser 65 70	
<210 > 25 <211 > 296 <212 > DNA <213 > H.Sapiens	
<4000 25 pandangatra télapaggak dadhasanga gametgagan engagatast gagantgata - 60	0
Equippings a goodcootecty eggengaged eegagtyget econopologie ggegagegeg ~ 12	0
octyacyctt coggggget gogoogstgo otgooscogg goottgatgg gagottcage 180	0
ggotoggago gotoatogoo ocaqogooo qqqotqqaca coaqoqqooto cacaqooaqo 24	0
ceeggt 24	6
<210> 26 <211> 80 <212> PAT <213> H.Sapiens	

Page 16

BNSC1901 - WCI - 196473A2T 14

<400> 26

Asn Pro Ile Ile Tyr Thr Leu Thr Asn Arg Asp Leu Arg His Ala Leu Leu Arg Leu Val Cys Cys Gly Arg His Ser Cys Gly Arg Asp Pro Ser Sor Gin Gin Sur Ala Ser Ala Ala Giu Ala Ser Gly Gly Leo Arq Arg Cys Leu Pro Pro Gly Leu Asp Gly Ser Phe Ser Gly Ser Glu Arg Ser Ser Pro Glm Arg Asp Gly Leu Asp Thr Ser Gly Ser Thr Gly Ser Fro Gly < 21.05 27 <211> 420 <212> DMP. <213> H.Sapiena <.220> <223> misc feature <222> $(81)^{-}$. (106) ₹223> n is any nucleic acid

400> 27
 cotqaaqaan aqoqocacca tgaccageat gtgcaccacg cocgctctgc gccgcgatec 60
 tegegggtee geagecteet nannnanna nannnanna nannnatge agagettgeg 120
 cgcgatgegg gegtacatga coacqatgaq egccagegge gccaggtaga tgtgcagagaa 180
 cagcacagtg gtgtagacce tgcgcatgce ctlotoggec caggecteec ageagagata 240
 cagagggtag gagtacatga gggcgtccac catgaagteg tgclcctcac gggtgacggt 360
 cagcgtgacg gccgagggac acatgatgaq cagcgccagg gcccagatga cggcgatggt 360
 cacqagcgcc ttccgcaggg tcagcttctc gcggaaaggg tccacgatga agoggaacct 420

```
<213> 139
<212> PRT
<23.3> H.Sapiens
<220>
<221> UNSURE
<222> (104)..(113)
<223> Xas is Unknown
```

29

<400> 28

<210>

Phe Arg Cys lle Vet His Pro Phe Arg Glu Lys Leu Thr Leu Arg Lys Page 17

1				5					10					15		
Ala	Leu	Val	Thr 20	lle	Ala	Val	Ile	Trp 25	Ala	Г∈л	Ala	Lau	Leo 30	Ile	Met	
Суз	Pro	Ser 35	Ala	Val	Thr	Leu	Th (*) 4 0	vəl	Thr	Arg	Glu	G).0 45	มาร	nis	Phe	
Mot	Wal 50	Азр	Als	Arg	Aen	Arg 55	Ser	Tyr	Fro	Leu	Tyr 60	Ser	Cys	Trp	Glu	
A19 65	Trp	Pro	G10	ьув	Gly 70	Met	AIG	Arg	Val	Тух 75	rat	Thr	Val	Leu	the BO	
Ser	Bie	rle	Tyr	Leu 85	Ala	Pro	Leo	Aln	Lav 90	Ile	Va)	Val	Met	Тут 95	Ala	
Azg	Ιle	Ala	Arg 200	Lys	Leu	Сув	Xea	Хаа 105	Хаа	Xaa	Xaa	Хаа	Хаа 110	Хаз	Xaa	
Xāā	Glu	Ala 115	Ala	Ast,	Pro	Arg	Ala 120	Ser	Arg	Arg	Arg	Ala 125	Arg	Val	Val	
ыів	Met 130		Val	Met	Val	Ala 135	Leo	Phs	Phe	ሞክェ						
<21 <21 <21 <21	1 -	29 318 DNA E.Sa	pien	ន												
<40 gen	0> 69 9 4	29 g og	tgag	test	ca g	geac	ttat	t ga	ggto	ettg	ttg	agea	gga	agea	gacaat	60
															գեզցցց	120
															осадац	180
															cttgsc	246
															рворься	300
		gge														318
<21 <20 <20 <21	.1≥ .1>	30 100 PRT H.Sa	ipien	ìΕ												
< d 0	0>	30														
Gly 1	y Pro	o Met	Pro	9 Pro	> Thi	Lei	a Lon	a (G1)	7 IV 10	e Ari	n Glo	s Ast	n 61:	у Ні. 15	s Ma	
Ala	a Sei	c Arg	9 Are 20	g Tæi	ı len	ն նդյ	y Mei	1 Asp 25	o Glo	Va)	l Ly:	s Glj	y 61° 30	o Ly	e Gln	
1.c	a G1;	y Ar	g Met	E Pho	е 'Гу:	r Al	a Il	e Thi	r Le	u Le Page		e le	u Le	u Le	и Тгр	

35 40 45

Ser Pro Tyr Ile Vol Ale Cys Tyr Trp Arg Val Phe Val Lys Ale Cys 50 60

 Δla Val Pro His Arg Tyr Leo Ala Thr Ala Val Trp Met Ser Phe Ala 65 -76 -75 -80

Gln Ala Ala Val Asn Pro fle Val Cys Phe Leu Leu Asn Lys Asp Leu 85 90 95

Lys Lys Cys Leo Arg Thr His Ale Pro Cys 100 0.05

<210> 31

<211.> 354

<212> DNA

<2130 H.Sapiena

<400> 31

tattotgtaa tgaagaatgt catloquat goosttggca catcoagteg cotcacctag 60 cattgtgaaa goostlogg. tggtgtattg coacttostt ttaasaggat goscaagtee 120 ctggtgcott tocacageas tgcsagtcat aggaggstt tetgtcacas cageggtags 180 ctggacaaat ggcaccatct tgcsaatgaa agoscotgca gtaaggaaat aggataaate 240 wlacatcasa acsaasagaa taaaggtto atotgtgtot ttgtaattat cactatoagt 300

costictgag octotgopaa aawgtitgat eeilgleett enichgtege cace 354

<210> 32

<211> 117

4212> PRT

ວ213> H.Sapiens

<400 > 32

Val Tyr Arg val lie Thr Ile Ile Lys Leo Phe Gly Arg Gly Ser Glo 1 5 10 15

Phe Wal Leu Met Tyr Asp Leu Ser Tyr Phe Leu Thr Ala Gly Ala Phe 35 40 45

lle Cys Lys Met Val Pro Phs Val Glo Sec The Als Val Vel The Glo 50 60

lle Leu Thr Mot Thr Cys lle Ala Val Glu Arg His Gln Gly Leu Val

Ris Pro Phe Lys Met Lys Trp Gln Tyr Thr Asn Arg Arg Ala Pho Thr 85 99

Met Leu Gly Glu Ala Thr Gly Cys Ala Asn Gly Ser Val Asn Asp ile Page 19

480

110 105 100 Leo His Tyr Arg Ile 115 <210> 33 <211> 621 <212><213> DNA H.Sapiens <4.00% gandaacatq otolttiitga agtacttgad ggtgtogtto tigaoggtoa dgaagdadag 60 aqtgttgako atgotottgo toatggogat gcactogacg atgtagaagg cagtgaggta 120 gtyctictoc iłoscabaca cygigyggaa gaagiogogo acgaigyiga agcociagaa 180 gggogocoag catagoacgt aggoggtgag gutqoacatg agcacoxqqo coqtettoot 240 gorgougogo aqeatotige qqatotgeld tytotgeaat edagggaddg collgaanda 300ყოლესდილეფ ფოლახიინუფ catagoacag ggtcatggtg accaegggge cosegaatte 360 tatgucsasg stasagsgga agtaggactt gtagtagage tgetggteca caggecagat 420

gtaggoggan gggatggoga toaggatgga caccgtocac accaaggosa toaggocagl 540 ggotqtttgg cacltoatto gtogtotoag oggatggada atagcoagat acctagggos 600

ctggoogcag sagatottit octggotott gucaatgacg aggacogtot ongtogtgaa

agaaqacaag tggaggcagc c 621

<210 · 34 <211 > 207

<212> PET

<213> H.Sapiens

<400> 34

Gly Cvs Loo Ris Leu Cys Ser Cys Pro Arg Tyr Leu Ala Ile Val Dia $\chi = 10$

Pro Leu Arg Pro Arg Met Lya Cya Gln Thr Ala Thr Gly Leu Ile Ala 20 25 30

Leo Val Trp Tor Val Ser Ile Leo Ile Ala Ile Pro Ser Mia Tyr Phe

Thr Thr Glu Tar Val Leu Val Ile Val Lye Ser Gla Glu Lya Ile Phe

Cys Gly Glm Ile Trp Pro Val Asp Glm Clm Lew Tyr Tyr Lys Ser Tyr 20 75 80

The Leu Phe II: 9he Gly II: Glu Phe Val Gly Pro Val Val Thr Met 85 90 95

The Leo Cys Tyr Ala Arg Ile Ser Arg Glu Leo Tep Phe Sys Ala Val	
Pro Gly Phe Glm Thr Glu Glm Ile Arg Lys Arg Leo Arg Cys Arg Arg 115 120 125	
Dys Thr Val Leu Val Leu Met Cys Ile Leu Thr Ala Tyr Val Leu Cys 130 135 140	
Trp Ala Pro Phe Tyr Gly Phe Thr Ilc Val Arg Asp Phe Phe Pro Thr 145 150 155 160	
Val Pho Val Lys Glu Lys His Tyr Leu Thr Ala Pho Tyr Ho Val Glu 165 170 175	
Cys lle Als Met Ser Asn Ser Met Ile Asn Thr Leu Cys Phe Val Thr 180 185 190	
Val Lya Asn Aap Thr Val Lya Tyr Phe Lys Lya Ile Met Leu Leu 195 200 205	
<210> 35 <211> 483 <212> DNA <213> O.Sapiens	
<400> 35 cagodacact geagtgatga astexaatgt ocaacaccaa coatagtoac cattactaac	60
teagraques caesauttos ettocagget hittosposge spagnesgag cocaggacag	120
genacacatg anagtitgada ggittelkeg geagnaecan cambaccana tampgoomgcam	180
gacagacagg cagcactdag tartgatgge arteageatg cteagecela ceaggtagge	240
aaaggteate aegetggtga agaagetagg gaaattgatg gagatggaac agaagaagtt	300
actgaggtac accaggeaat ttataatetg gaagcagagg aagaggaagt eggecneepn	360
caggetgagg wegtagaeng ugaaggegtt eetgegeatg eggaageesa ggagesagag	420
cacaeachig bhtoclacha godinacces ggcaabgeas angatoagga agacogggat	480
cag	483
<210> 36 <211> 161 <212> PRT <213> H.Sapiens	
<400> 36	
Leo The Pro Val Phe Leo He Leo Phe He Ala Leo Val Gly Leo Val 1 5 10 15	
Gly Asn Gly Phe Val Leo Trp Leo Leo Gly Phe Arg Met Arg Arg Asn 20 25 30	

Ala Fhe Ser Val Tyr Val Leu Ser Leu Ala Gly Ala Asp Phe Leu Phe $35 \hspace{1cm} 40 \hspace{1cm} 45$												
Leu Cya Phe Gln Ile Ila Asn Cys Leu Val Tyr Leu Ser Asn Phe Phe 50 55												
Cys Ser Ils Ser Ilc Asn Phe Pro Ser Phe Phe Thr Ser Val Met Thr 65 70 75 80												
rhe Ala Tyr Lau Val Gly Leu Ser Met Lou Ser Ala Ile Ser Thr Glu 65 90 95												
Cys Cys Leu Ser Val Leu Arg Fro Ile Trp Tyr Cys Cys Cys Cys Pro 100 105 110												
Arg Asn Leu Ser Thr Val Met Cys Ala Leu Pro Trp Ala Leu Ser Leo 115 120 125												
Leo Leo Asm Thr Loo Glo Gly Lys Pho Cys Cly Phe Leo Val Ser Asm 130 135 140												
Gly Asp Tyr Gly Trp Cys Trp Thr Phe Asp Phe Ile Thr Ala Val Trp 145 150 155 160												
1eu												
<pre><210> 37 <211> 530 <212> INA <213> H. Sapiens</pre>												
-400> 37 gagagtotga ttotgactta cateacatat gtaggcotgg geatterat ttgcsycotg 68												
etecttiget igteogitga ggiteetagie iggageeaag igacaaagae agagateaee 1												
tatttacces atgigigeat igitaacatt geagseacit igeigaigge agaigigidg — 180												
Theathylog officettet Laglygonea ataacacac acaagggatg tgttgcagec 240												
adattititing of cattrett thacetter glabbithet goodgetige eargeacte 300												
cttatcctct atggastcst gattgttttc 330												
<pre>:210> 38</pre>												
<400> 38												
Glu Ser Leu Ile Leu Thr Tyr Ile Thr Tyr Val Gly Leu Gly Ile Ger 1 5 10 15												
lle Cys Ser Loα Ile Lou Cys Loα Ser Val Clu Val Leu Val Trp Ser 20 25 30												
Page 22												

628

Gln Val Thr Lys Thr Glu Ilo Thr Tyr Leu Arg His Val Cys Ile Val 35

Asn Ile Als Als Thr Leu Leu Met Als Asp Val Trp Phe Ile Val Als Sp Che Leu Ser Gly Pro Ile Thr His His Lys Gly Cys Val Als Als Als Thr Fhe Che Gly His Phe Phe Tyr Leu Ser Val Phe Phe Trp Met Leu

Ala Lys Ala Loo hou lie too Tyr Giy lie Met lie Val Phe

<210> 39 <211> 628

1212> DNA 1213> B.Sapiens

WO 01/36473

<400> 39 tigiciqqes qiaqaqaqat qicaqqetto aqaqtessca agsaciqqat iicaaactqq 60 atttgaggae coccaecttt ggtaagtgae ttattatetg egageetetg titetetett 120 otifiaaatoa qqacaqbaaa toocatacqq caqqqtqqtq qqqaqaat<mark>ca qaqatqatac</mark> 160 240 agetygtgat cacatetygt tigtyticco aggegoacce gackaeggkt totgagostg gatebaabog toocagtott oggtacaaaa otgacaccaa tbaaboggsog tgaggagaek 300 cettgetaca atcogaccet gagetteacg gtgetgacgt geatcattte cettgtegga 360 ctoacaqqae aceqqoteek qolobqqoko olqqqotaco qoatqoqcaq qaaoqotqto 420 recatetare tentrascet ggoograges gantiquiel teetesgett coagattata 480 egitegedat taegeoteat caatatoage cateteated geaaaateek egibbetete 540 ateacettic ectactitae aggeotgagt atgetgageg coateageac ceayegelye 600

<2105 40 <211> 205 <212> FRT <213> H.Sapiens

Abstatighte intropectat etemplace

<400> 4D

Leu Cyn Cly Ser Arg Glu Met Ser Gly Phe Arg Val Asn Lys Asn Trp $1 \ 5 \ 10 \ 15$

fle Ser Aso Trp fle Giy Pro Pro Pro Lou Val Ser Asp Leu Leu Ser 20 25 30

Ala Ser Leu Cys Phe Ser Leu Leu Met Arg Thr Val Asn Pro Ile Arg 35 40 45

Gln	61y 50	Gly	G1 y	610	Asn	61n 55	Arg	Туг	Ser	Trp	Ser 60	His	Leu	Val	Сув	
Val 65	Pro	Arg	Ģ1y	Thr	Arg 70	Leu	Gly	Fhe	Leu	ser 75	Met	Asp	Pro	ТПГ	Va1 80	
Pro	Val	Phe	Gly	Thr 85	Lys	Leu	Thr	Pro	11e 90	Asn	Gly	Arg	Glu	Glu 95	ፐ ንኮ ሮ	
Pro	Сув	Туг	Asn 100	Gln	Thr	Leu	Ser	Phe 105	Thr	Val	Leu	Thr	Сув 110	Ile	Ile	
Ser	Leu	Va) 13.5	61 y	Leu	Thr	Gl y	Asn 130	Ale	Val	Va.1	Lou	Trp 125	Leu	Leu	Gly	
Tyr	Arç 130	Met	Arg	Arg	Aen	Als 135	Val	Ser	Ile	туг	11e 140	Io⊕ o	ħsn	Leu	Ala	
Ala 145	Ala	Asp	Phe	ren	Phe 150	ren	Ser	[·he	Gln	11e 155	Ile	Arg	Ser	Pro	Leo 160	
Yrd	Lau:	Ile	Asn	I10 165	Ser	Ris	Leu	Ile	Arg 170	Lys	Ile	Leu	Val	Ser 175	Val	
међ	The	Phe	Pro 180	туr	Phe	Thr	Gly	185	Sec	Met	Leu	Sor	Ala 190	Ile	Ser	
Thr	Glu	Arg 195	Сув	peo	Ser	val	Leu 200	Trp	Pro	Ile	Trp	7yr 205				
<21 <21 <21 <21	1> 2>	41 319 DNA H.Sa	p) en	s												
<40 aca	0> gaaa	41 gca	ತಿ ರ ್ಕ್ಷರ	cacc	ag g	acot	tagg	c at	agto	atgg	gag	tgt:	tgt	ցենց	tgotgg	60
															gaayat	120
															ttatat	180
												240				
												300				
caataggact cttttctgg											319					
<21 <21 <21 <21	1= 2>	42 103 PRT H.Sa	apier	នេ												
<40		12											_		F-1	
Thi 1	Glu	Sei	г Буа	: Ala 5	Thi	: Arq	Thi	Let	10 10	/ Ile	e Val	. Met	: Gly	/ Va!	Pho	
	Page 24															

Val Leo Cys Trp Leo Pro Phe Pha Val Leo Thr Ile Thr Asp Pro Phe lls Asn Phe Thr Thr Leo Glu Asp Leo Tyr Asn Val Phe Leo Trp Leo Gly Tyr Phe Ash Ser Als Phe Ash Pro Ile Leu Tyr Gly Met Leu Tyr Pro Trp Phe Arg Lys Ala Leu Arg Met Ile Val Thr Gly Met Ile Fhe His Pro Asp Ser Sor The Leu Sor Leu Phe Ser Ala Ris Ala Ala Val Phe Ile Ile Glm Asp Ser Phe 43 <210> 515 <211> <212> DNA <213> B.Sapiens <40ft> 43 taggaatoto agagaagawa gtuaggaaco agawaaccat awaagaatgi abatggaaas 60 qualvegosa stottetica ottetoacte satotessal algtomasat acatgaagad 120 180 sangaatoot ttaggagaad tottosatot attotoctad aacttoocat algahoolige ttgcotetet atgtccaagt gittattitt geagttgace ttaatticaa gitagittig 240 aggiototae agizatgitt itaatoigio totactioti cagaaaataa attagiigti 300 egendatosq tootlaaqad olligoogobb acaabaagtt tiattgeett eecaaaceat 360 tyytaaaaga asycataaat casyggytto atagolysat kalaataaan acaccamact 420 assatotoat associasog aggaqttata assitoatat asgostosat cactgostoa 4 80 515 acqaqqtatq qtaqccaaqa qacaaqaaat qctqc

<210> 44

<211> 148

<232> PRT

<213> IL Sapiena

<400> 44

Leu Ris Gl
n Arg Gly Met Val Ala Lys Arg Gl
n Glu Met Leu Ala Ala 1 $$ 5 $$ 10 $$ 15

Pho Leo Val Sar Txp Leo Pro Tyr Leo Val Asp Ala Val Ile Asp Ala 20 25 30

Tyr Met Asn Phe lie Thr Pro Pro Tyr Val Tyr Glo Jio Leo Val Trp 35 - 40

Cys Vel Tyr Tyr Asn Ser Ala Met Aan Pro Leu Ile Tyr Ala Phe Phe 50 60	
Tyr Gin Trp Pho Gly Lys Ala He Lys Leu Hio Val Ser Giy Lys Val 65 70 75 80	
Leu Ang Thr Asp Sen Ser Thr Thr Amn Leu Pho Sen Glu Glu Val Glu 85 90 95	
Thr Asp Lya His Tyr Cys Arg Asp Leu Lys Thr Asn Leu Lys Leu Arg 100 105 110	
Ser Thr Alm Lys Ilo Asn Thr Trp Thr Arg Gly Lys His Asp His Met 115 120 115	•
Pro Ser Cys Arg Th: The His Ser Thr Val Val Leu Lys His Leu Leu 130 135 140	
Ser Ser Cym Ile 145	
<210: 45 <211: 726 <212: DNA <213: H.Sapiens	
<40005 45	£ 0
ntggaaayag iteetegate tateetetae geogteettig gittitiggige tigtigetigges	60
gegittiggas acttactigst catgatteet atcetteact tetascaset geacacacet	120
acaasottto igatigogio poigpoolgi gotgactici iggigpgagi caciqigatg	180
coottoagoa cagteaggto tetuesques byfttggtact ttgggesoug ttactgtada	240
$the catacat \ \textit{oblitte} \textit{parable} \ \textit{abgolightet}. \ \textit{this golithet} \ \textit{this parable} \ \textit{abgolightet}.$	300
tototigata gatacattgc tgttactgat cototgacct atocaaccaa gtttactgtg	360
teagttteag ggatatgeat tgttetttee tggttetttt etgteaeata cagetttteg	420
atottitada ogggogodaa ognagaayga attgaggaat tagtaybbgo botmacolgb	4(4)
gtangangunt eccaegeton annuasteas asciggiten tactitigiti tettetante	540
Uttabannoa atgtogocat götgittata tabagtaaga tattitiggi ggocaageat	600
caggotagga agatagaaag tacagocago caagotcagt cottotcaga gagttacaag	660
gaaagagtag caaaaagaga gagaaagget goosaaaect tgggaattgo tatggcogoo	720
tttett	726

Page 26

<210> 46 <211> 241 <212> PRT <213> H.Sapiens

<400> 46

Len Glu Arg Gly Pro Arg Ser Ile Len Tyr Ale Val Len Gly Phe Gly 1 5 10

Als Val Leu Ala Ala Phe Gly Asn Leu Leu Val Met Ile Ala Ile Leu $20 \hspace{1cm} 25 \hspace{1cm} 30$

His Phe Gln Leu His Thr Pro Thr Asn Pho Leu Ile Als Ser Leu Ala 35 40 45

Cys Ale Asp Phe Leu Val Gly Val Thr Val Met Pro Phe Sec Thr Val 50 60

arg Ser val Glu Ser Cys Trp Tyr Pbe Gly Asp Ser Tyr Cys Lys Pbe G5 -70 -75 -80

His Thr Cys Phe Asp Thr Ser Phe Cys Phe Ala Ser Leu Phe Rie Leu 85 90 95

Cys Cys Ile Ser Val Asp Arg Tyr Ile Ala Val Thr Asp Pro Leu Thr 100 105 110

Tyr Fro Thr Lys Phe Thr Val Ser Val Ser Gly Ile Cys Ile Val Leu 115 120 125

Ser Trp the the Ser Val Thr Tyr Ser Phe Ser Tle Phe Tyr Thr Gly $130\,$

Ala Ash Glu Glu Gly Ile Glu Glu Leu Val Val Ala Leu Thr Cys val 145 155 160

Gly Cly Cys Gln Ala Pro Leu Aso Glo Aon Trp Vel Leu Leu Cys Phe 165 170 175

Leu Leu Fhe Phe Ile Pro Asn Val Ala Met Val Phe Ile Tyr Ser Lys 180 180 180

lle Phe Leu Val Ala Lys His Gln Ala Arg Lys Ile Glu Ser Tbr Ala 195 - 200 - 205

Ser Gin Ala Gla Ser Phe Ser Glo Ser Tyc Lys Glu Arg Val Ala Lys 210 215 220

 Δrg Glu Δrg Lys Ala Ala Lys Thr Leu Gly IIe Ala Met, Ala Ala Pho 225 230 235

Leu

<210> 47

<211> 660

<212> DNA

<213> H.Sapiens

<400> 47

saccegging contacted aspecded goottgicts typectitat cascageigt

otosatopag	thatateligt	cttcattggg	catgacttct.	gggagdacht	getecastes	120
otgotaposo	ccttagescg	ggcacttagc	gaggagadag	alagtgootg	aatoccaget	180
сссадусаца	tesyteettt	ataacatgac	ccaatttcct	actocatitt	Cocaccasts	240
Batcotette	ccaaacaget	ctaccatast	ccaacateca	acagaattta	а <u>о</u> вцаараа	300
ccacaacttt	taagtgaget	ctatgtgcta	ggtcatgttt	tegsatscas	ccttaagl;qc	360
otggwagatg	дардомадма	асдаясаедд	totoattott	tagaggaaga	cagticacca	420
agacticaea);	ацанананар	abaghtatot	toligaceeaa	cungteataa	aattgggtca	180
ggaeetçdag	caatgacttt	atgctagaat	ccagageact	agcaggaaac	ipottaeett	540
ttacttaatc	saagtcaagt	ttggacatsc	atgtcaggta	aaac ctag ca	gagatgsgot.	600
accttgattt	taaaacttea	agggataget	caathteate	aagatoottt	tgatgacttg	660

<210> 48

<400> 48

Asn Glm Val Ala Leu Leu Leu Arg Pro Leu Ala Leu Ser Met Ala 2he

Hic Asm Ser Cys Len Asm Pro Val Leu Tyr Val Phe Hie Gly His Asp

Phe Trp Giu Nis Leu Leu Nis Ser Leu Leu Ala Ala Leu Giu Arg Ala

Leu Ser Glu Glu Pro Amp Ser Ala lle Pro Ala Pro Arg Glo Met. Sen

Pro Leu His Asp Pro Ile Ser Tyr Sor Ile Phe Pro Pro Leu Asn Pro

Len Pro Lys Gin Leu Tyr His Ash Pro The Ser Ash Arg Ils Gle Ash

bys Pro Gln Leu Leu Ser Glw Leu Tyr Val Leu Gly His Val Leu Glo

Tyr Asn Leu Lys Cym Leu Glo Asp Gly Gly Lys Lys Sin Thr Arg Ser 115 120 125

Die Ser Leu Glu Glo Asp Ser Ser Pro Arg Leu Lys Gla Lys Lys Arg

Leu Ser Cys Asp Lys Thr Ser His Lys Ile Gly Ser Gly Pro Ala Ala

Met Thr Len Cys Ash Pro Glu Bis Gln Glu Thr Ala Ile Leu Leu Ash Page 28

<211> 211 <212> PRT

<213> 8 Sapiene

165 170 175

Glo Ser Glo Val Trp Thr Tyr Met Ser Gly Lya Thr Glo Arg Ala Thr

Led He Leu Lys Leu Clm Gly Ho Ale Glm Cys Ris Glm Asp Pro Phe 200

Asp Asp Leu 210

<210> 48 <211> 465

<212> DNA <213> H.Sapiens

<480> 49

ecttotteae moceancale oleaagetet leograceea egaqeequae qqoeqqeaqu 60 120 aqooqaqqqq qqooqtqqqq etqqcoqqqq tqqtcttqct qqeetthqtc accbqcbtcq copocaseas officiales of engagesca togigações congitates engagaget 180 actaceacqt qtaceacctc acqctqtqtc tcaqctqcct caacaactqt ctqqacccqt 240 ttgtttatta etlitguatee egggaslitee agelgegeel gegggaalat tigggetgee 300 gopyragingo pagagarang piggarango operanga at totototototo operangu 360 nquanqtong etergagger ggtgregearr etgsagggat ggagggagri accapgoung 420 465

c230> 50

:231> 160

<212> PRT

:81**13**> B. Sepions

<40n> 50

ben Phe Thr Ala Thr Ile Leu Lye Leu Leu Arg Thr Glo Glo Ala Dis

qeeteeagag qeaggagagt gtgttetgag teeeggggge geage

Gly Arg Glu Glm Arg Arg Arg Ala Val Cly Leu Ala Ala Val Val Leu

led Ala Pho Val Thr Cys Pha Ala Pro Ash Ash Pho Val Led Led Ala

His lle Val Ser Arg Leu Phe Tyr Gly Lya Ser Tyr Tyr His Val Tyr

Lys Leu Thr Leu Cys Lou Ser Cys Lou Ash Ash Cys Lou Asp Pro Phe

Val Tyr Tyr Phe Ala Ser Arg Glu Phe Gln Leu Arg Leu Arg Glu Tyr

Lou Gly Cys Arg Arg Vel Pro Arg Asp Thr Lou Asp Thr Arg Arg Glo 100 105 110	
Sec Leu Pho Sec Ala Arq Thr Thr Ser Val Arg Ser Glu Ala Gly Ala 115 120)25	
His Pro Glu Gly Met Glu Gly Ala Thr Arg Pro Gly Leo Glo Arg Glo 130 135 140	
Glu Ser Val Phe Val Pro Gly Ala Gln Ala Ala Pro Pro Gly Leu Arg 145 150 155 160	
<210> 51 <211> 503 <212> ONA <213> H.Sapiene	
<4000 51 ttacttatto tgoccittat coaacttita attoccittg ciattotect goctoatti.	60
etymaeteat titeeetaft ateetgeete acattgatea agggatgagg etggeaggat	120
orquaectus cappycoorg typpocatga paygotholy pactigeaco torgazozot	180
conscious officeades destablished factories advantages and advantages of the conscious of t	240
gtggagagee ataggetatt ggggtggaea ggettgggtg ceteatggga geteeccatg	300
ggaçotgtmg cotottgggm cotottattt otdacococag gotttococky gagagettma	360
agteagaaga tgoccoaaag atocaagtag coatgagtag cagootatto etectgaate	420
tequestett egstevaleta gagaajlagust cavagegete igaiçetgee igeigggese	480
gyggggotyt ettecantae tteetgetet gtgeellead elgganggge ottgaageet	540
tecaceteta petgeteget gteagggtet teaacaceta ettegogeae Lacttophys	600
age	603
<210> 52 C11> 198 <12> PRT <113> B.Sapiens	
<400> 52	
Glu Thr Tyr Ser Ala Leu Tyr Pro Thr Phe Asn Ser Leu Cys Tyr Ser 1 10 15	
Pro Ala Ser Pho Sor Gly Lou Ile Phe Pro Ilo Ilo Lou Pro His Ile 20 25 30	
Asp Gin Gly Met Arg Leu Ala Gly Ser Gly Thr His Arg Ala Pro Trp $$35$$ $$40$$ $$45$	
Als Met Arg Gly Ser Trp Thr Thr Ser Gly His Ser His Ser Gly Cys 50 50 50	

Arg Glm Gly Trp Lys Leo Asp Gla Glo Ala Gly Ala Gly Ser Gly Gly Gly Glo Pro Ala Tle Gly Val Asp Arg Leo Gly Cys Neu Met Gly Ala Pro His Gly Ser Cys Gly Pro Leu Gly Pro Leu Ile Ser His Pro Arq Leo Ser Arg Glu Arg Phe Lys Ser Glu Asp Ala Pro Lys fle His Val Ala Leu Gly Cly Ser Lau Pho Leu Leu Asn Leu Ala Phe Leu Val Asn vel Gly Ser Gly Ser Lys Gly Ser Amp Ala Ala Cys Trp Ala Arg Cly Ala Val Phe His Tyr Phe Leu Leu Cys Ala Phe Thr Trp Met Gly Leu Clu Ala Phe His Lou Tyr Lou Leu Ala Val Arg Val Phe Asn Thr Tyr Phe Gly His Tyr Phe Leu 195 <210> 53 <211> 395 DWA. <21**2**> <213> H.Sapiens <4000 saltgytogy agagtgesge tgettgaaat ggaggattga aabeatcace aggaggiitte 60 caaacacago cagcacagoo ocaaagocaa acactatgta cagaatcaco coggatocog 10.0 goqaqaaqqq qattttoaca caggadocat toacqttoqo qtagoacago tqoacagooa 1 23 11 ccagcaggga tgaattgetg etcataæege tggtatttae atatggagaa attitgteet 240 tgttgattat dacaasaaat abaggattyt tootgatttt dattgotoot yoggasaaaa 300 335 acacatatte accaggatge cagaggaaat gatou <210> 54 111 <211>

<211> 111 <212> PRT <213> B.Sapiens

<400> 54

App Ris Phe Leo Trp His Pro Gly Glo Tyr Val Phe Phe Ser Ala Gly 1 5 10 15

Ala Met Lys fle Arg Asn Asn Pro Val Phe Phe Val Ile 11e Asn Lys 20 - 25

Asp Lys Ile Ser Pro Tyr Val Asm Thr Ser Val Met Ser Ser Asm Ser Ser Lou Lau Val Ala Val Gin Leu Cys Tyr Ala Asn Val Asn Gly Ser Cys Val. lys lite Pro Phe Ser Pro Gly Ser Arg Val Lie Leo Tyr Ilo Val Phe Gly The Gly Ala Val Leu Ala Val Phe Gly Ash Leu Leu Val Met Ile Ser Ile Leu Ris Phe Lys Cln Leu His Ser Pro Thr Ash <210> 55 <311.5 586 <2125 DNA <213> R.Sapiene <400> 55 cacatottaa caaqaotqaa aaacattoat ttqtttttaa tttgaaqaqo aatttatttq 60 ctatteatte ataqtettne ttqattttta aaaacteatt tegettgeta attttaaage 120 tatoetgame tinglotein cascigotta labatettes gasseconet testigotte 180 tgasciette ittassacri gaccagitze asisaniitti sligettiec tessecustyn 240 qtaxaatsaa qoataaatoa aaqqattoat qqotqaqtta tsataaqoac accascagoa 300 toataaatac aggoaggggt tatwaagcoo ataaaggoat caattaatga atcaatgota 360 420 tatggtaacc atreastcat asobyetsee schutgacce ccaggglitti agotgctitti etebatetee lyneeachet gyettligtaa etetetgagg algattetek etkoekaesa 460

<210> 56 <211> 190

<212> PRT

<213> H.Sapiens

<400> 56

Len Thr Asp Phe Len Ser Phe Phe Ile Pro Thr Phe Ile Met Ile Ile

glabbitkopa tottilitogo etgiogicia gecacazgaa ataigitace ataesyaatt

atoataataa aggtaggtat aaagaaggat agaaaatotg toaaca

Leu Tyr Gly Asn Ile Phe Leu Val Ala Arg Arg Gin Ala Lys Lys Ilo 20 25 30

Glu Asn Thr Gly Ser Lys Thr Glu Ser Ser Ser Glo Ser Tyr Lys Als 35 40 45

Page 32

540

586

Arg Val Als Arg Arg Glu Arg Lys Als Als Lys Thr Leu Gly Val Thr 50 60	
Val Val Ala Phe Met Ile Ser Trp Leu Pro Tyr Ser Ile Asp Ser Leu 65 70 75 80	
The Asp Ale Phe Met Gly Phe The The Pro Ale Cys The Tyr Glu He 85 90 95	
Cys Cys Trp Cye Als Tyr Tyr Asn Ser Ala Met Asn Pro Leu Ile Tyr 100 105 110	
Ala Leo Phe Tyr Fro Trp Phe Arg Lys Ala Ile Lys Val Ile Val Thr 115 120 125	
Gly Gln Val Leu Lys Asn Ser Ser Ala Thr Met Asn Leu Phe Ser Glu 130 135 140	
His The Ale Val Gly Thr Lys Pho Arg The Pro Lee Lys Lee Pro Ser 145 150 255 160	
Glu Met Ser Phe Lys Ser Ser Lys Thr Met Aen Glo Gin Ile Aen Cys 165 170 175	
Ser Ser Asn Lys Gln Ile Asn Val Phe Gln Ser Cys Asp Val 180 185 190	
<210> 57 <211> 976 <212> DNA	
<213> 0. Sapiene	
<400> 57 tttgtggcaa ggagaccetg atceeggtet teetgatest titcattgcc etggteggge	60
tggtaggaaa ogggtttgtg clologotoc teggetteen catgegeamy aargeettst	120
engtolaegt deteagonly groppygody acttentett entetgetic magabbabaa	180
stigocologi glaceleagt aactiettet giteratete cateaattic metagette(24 D
teaccartgt gatgacetgt geotacettg caggeotgag catgotgage accgteagea	300
cogagogoty cotytocyto otytogodda totygtatog otygogoddo oddadadaec	360
tgtcagcggt cgtgtgtgtc ctgctctggg ccctgtccct actgctgacc atcttggaag	420

ggaagttetg tggettetta titaglyalo qtgaetetqg liegteteag acattigati

teateachge agegtegetg attittitat teatggitet eigigggien agietygeen

byckgytowy gatectetyt gyetecaggy ytetgecast gaccaggety tacetgacca

tectgetese sgtgetggtg teceteetet geggeetgee etttggeatt cagtggttee

taatattate gatotggaag gattotgate totlattite lootatioal coogittoag togtootgto atotoltaan agosgigoos accommist tiantiotic gigggetoil

Page 33

480

540

€00

660 720

780

ttagga	agea gtqqceqstq	cagdaccoga	tocleasget	ggetetecaq	agggelet g e	840
ა ცემიგ	ttigo tgaggtiggat	cacegtgaeg	gstgetteeg	teagggeace	იფფოდასსია	900
аадаад	catt ctggtgtagg	gatggacccc	totacttoca	tcatatatat	gtggetttgs	960
daddes	actt tgeece					976
<2105 <2115 <2125 <2135 <2205 <2215 <2225 <2235	PRT					
<400.>	58					

Cys Gly bys Glu Thr Leu Ile Pro Val Phe Leu Ile Leu Phe Ile Als 1 $^{\circ}$

Leo Val Giy Leo Val Gly Ash Gly Phe Val Leo Trp Leo Leo Gly Phe $20 \\ 25 \\ 30$

Arg Met Arg Arg Asn Ala Phe Ser Val Tyr Val Leu Ser Leu Ala Gly 35 40 45

Ala Asp Phe Leo Phe Leo Cys Phe Glo Ile Ile Aso Cys Leo Vel Tyr 50 - 60

Leu Ser Asn Phe Phe Cye Ser Ile Ser Ile Asn The Pro Ser Phe Phe 65 70 75 80

Thr Thr Val Met Thr Cys Als Tyr Loe Als Gly Leu Ser Met Leu Ser 85 90 95

Thi Val Ser Thr Glu Arg Cys Leo Ser Val Leo Trp Pro Tie Trp Tyr 100 105 110

Arg Cys Arg Arg Pro Arg Bis Leu Ser Ala Val Val Cys Val Leu Leu 115 120 125

Trp Ala Leu Ser Leu Leu Ser Ile Leu Glu Gly Lys Phc Cys Gly 130 140

Phe Leu Pho Sar Asp Gly Asp Ser Gly Trp Cys Gln Thr Phe Asp Phe 145 155 160

Ile Thr Ala Ala Trp Leo Ile Phe Leo Phe Met Val Leo Cye Gly Ser 165 170 170

Ser Len Ala Leu Leu Val Arg Ile Leu Cys Gly Ser Arg Gly Leu Pro-180 135 190

Leo Thr Ang Leo Tyr Leo Thr Ilo Leo Leo Thr Vol Leo Val Ser Leo Page 34

BNSCHAD KWT HE475AZT

195	200	205
Leu Cys Gly Leu Pro Phe Gl 210 21		
Trp Lys Asp Ser Asp Val Le 225 230	α Phe Cys His Ile Hi 235	is Pro Val Ser Val 240
Val Lau For Ser Leu Ash Se 245	r Ser Ala Aen Pro II 250	le Ile Tyr Phe Phe 255
Val Gly Ser Phe Arg Lya Gl 260	n Trp Arg Xaa Glo Hi 268	is Pro 11e Leu Lys 270
Leu Ala Leu Glm Arg Ala Le 275	u Gln Asp Ile Ala Gl 280	to Val Asp Bis Ser 285
Glu Gly Cya Pha Ara Glm GA 290		os Glo Ala Phe Trp 36
Cys Arm Asp Sly Pro Let Ty 305	r Fhe His His lle Ty 315	yr Val Ala Len Arg 320
Gly Asn Phe Ala		
<pre><210> 59 <211> 578 <212> DNA <213> E.Sopions</pre>		
<4085 59 htttgdetct cactgttgag caga	caycot qotqaasqtt qt	ogotgaco accadatata 60
gtaacaggtt accaaaggtg ttca	gaçcaş cataatgoto ta	nçamacçat gimaçetica — 120
tygatotgat totoaatgga acaa	otçatt galagçagça tç	gaņattoga toetņastga — 180
pootoaagat atggaagggt वस्त्र	aaceta cutaasetgo as	arqaqtaqo agaatqqtta 240
geoktomige blbetgetta agge	agetgt cagtttgeag to	rcatqqqto easqtqtqqs — 300
taatogtggt atagcasagt gtca	ctatoa cosaggggag go	cagasagts ettgcagtes 360
aaatcaggtt gtaccactta atag	tattga gitteateeça ad	otogtgagg togaçacagg 420
otgatotytt yyteebytty ytte	atobqa toaaqaaqgi ca	aloggaalg acametacea — 400
gtgaaatgat coacaccaca gcac	aggeta caactgeaca to	ngagtitty tgaatggasa — 540
ageagetest toggtgsate ates	cacagt agoggaag	578
<210> 60 <211> 292 <212> DRT <213> H.Sapiens		

Page 35

<400> 00

Pho Ang Tyr Cys Val Ile His Pro Met Ser Cys Phe Ser Ile His 1 10 15							
Lys Thr Arg Cys Ala Val Val Ala Cys Ala Val Val Trp Ile Ile Ser 20 25 30							
Len Val Ala Val Ils Pro Met The Fhe Leo Ile Thr Ser The Ash Apq 35 40 45							
The Ash Arg Ser Ala Cya Leu Asp Leu The Ser Ser Aap Glo Leu Ash 50 55 60							
Thr fle Lys Trp Tyr Asn Leu IIe Leo Thr Ala Ser Thr Phe Cys Leu 65 70 75 80							
Pro Leu Val Ile Mal Thr Leu Cys Tyr Thr Thr Ile Ile Mis Thr Leu ×5 90 95							
Thu Dis Giv Leu Glu Thr Asp Ser Cys Leu Lys Glu Lys Ala Arg Arg 100 105 110							
Leu Thr Ile Leu Leu Leu Ala Phe Tyr Val Cys Phe Leu Pro Phe 115 120 125							
His He Leu Arg Val He Glm App Arg He Ser Ala Cys Phe Glm Ser 130 135 140							
Val Val Pro Leu Arq Ile Arg Ser Met Evs Leo Thr Ser Phe Leo Asp 145 150 150 155 160							
His Tyr Ale Ale Leu Asn Thr Phe Gly Asn Leu Leu Leu Tyr Val Val 165 170 175							
Val Ser Asp Ash Phe Gln Gln Ala Val Cys Ser Thr Val Arg Cys Lys 180 185 190							
<210> 61 <211> 872 <212> DWA <213> H.Sapiene							
<400> 61 gggagggete gtagasacae taescetace etttetgttt etteeteete ttteetitee	60						
atetgittot catggiotec igleketete teletetete recinitiet eintechuge	120						
totttotcat doodbinatt telgtgtcaa totcaatees titatategg tggocaetit	180						
totatotott tyliciatot etetetetet etettieees etitytetet geaegeeigt	240						
totalittic typestates tetestiques teatestetet etestetet typesteate	300						
tenengiere tergigiere igteressee gereattees all'hosagal gematytone	360						
aggacaacto atggagoodo odogogooda togaqtadog gastqqctga odooctaggq	420						

Page 36

480

ttggcagtag cocctgacco tcombabago chacketaco ggagagactg aggagatgag

eggegetetg	topecapogt	cogoateage	ttatgtgaag	ctggtactgc	tgggactgat	540
tatgtgcgtg	agootagogg	gtaaugecat	cttetecetg	ctggtgstca	aggagogggo	600
cotgoacaag	getectiset	acttootget	ggacolgtgo	ctggccgatg	gestaegete	660
tecoglotes	ttoccctttg	tgetggette	tgtgcgccac	ggetetbeat	eganottoug	720
tgeseteage	tgcaagattg	tggcotttat	ggoogtaate	titigettae	atgaggeath	780
catgotgtto	tgcatcagcg	toaccegeta	catçgocato	godeaceacc	gettetaege	840
caagogoatg	acaetotoga	catgogoggo	tq			872

<20,0> 62

<211> 143

<212> PRT <213> H.Sapiens

<400: 62

Most Ala Asa Tar The Gly Giu Pro Glu Gio Val Ser Gly Ala Lou Ser 1 5 10 15

 ho_{10} ho_{20} ho_{20} Ata Ser Ala Tyr Val Lys Leo Val Lau Leo Gly Leo 100 25 30

Met Cys Val Ser Leu Ala Gly Asn Als lie Leu Ser Leu Leu Val Leu 35 40 45

bys Glo Arg Ala Leo His Lys Ala Pro Tyr Tyr Phe Leo Leo Asp Leo
50 55

Cys Lou Ale Asp Gly Tie Ang Ser Ale Val Cys Phe Pro Phe Val Leu 65 76 80

Ale Ser Val Arg Bis Gly Ser Ser Trp Thr Phe Ser Ala Leu Ser Cys 85 90 95

Lys The Val Als Phe Met Als Val Leo She Cys Phe His Als Als Phe $100 \hspace{1.5cm} 105 \hspace{1.5cm} 110$

Met Leo Phe Cys Ilo Sor Val Thr Arg Tyr Met Ala Ile Ala Ris Ris

Arg Phe Tyr Als Lye Arg Met Thr Leu Trp Thr Cys Ala Ala Glu 130 $$135\$

<210> 63

<211: 962

<212> DNA

<213: H.Sapiens

<400% 63

ammantiget gisetgaaci siigaatgga actiggsaat aaagteesti eeaaastaac 60 Labteiteas cagagagtaa taggtaaatg tiitagaagt gagaggaete amatigeema 120

securence cadedadean coddenness consumes, dadadamen

tgatttactc	tettattttt	cotoctaggt	ttotgogata	agtatgtgca	aataaaaaat	180
asacatgaga	aggaactgta	acobgattat	ggatttggga	aaaagataau	temacacaca	240
ងងកូចុំផ្ទុកកុសមួ	taasetgatb	gacageeete	aggmatgatg	coottitgou	amaatataat	300
lastatittco	tqtqtqaaea	acsactggtc	aaatgatgto	ogtgetteen	tgtacapttt	360
aatggtgete	ataattotga	ocacactogt	tggcaatctg	atagtbettg	tttctstate	480
acacttcass	caacttosta	0000630863	ttggctcatt	cattccatgg	ccactgtgga	480
ctttcttctg	gggtgtotgg	teatgootta	cogtatggtg	agatotgotg	agcactgttg	540
gtatüütgga	gaaglobbot	qtaamattop	nacaagnees	pacattatge	tqugeteage	600
checatittle	catttqtctt	tcatctccst	tgadegetse	tatgotylgt	ghyahimaet	660
gagatatasa	gecaagatga	atatottggt	tatttgtgtg	atqatettea	ttagttggag	720
tgtecetget	gtttttgcat	ttggaatgat	ctttctggag	ctanactica	aaggogotga	780
agagatatat	tacaancatg	ttpackgpag	aggaggttgc	tolytottot	ttugeasaat	940
atotggggts	otgacottta	tgachbothb	tlahetacct	ggalotatta	buttatņtūt	900
ctattacaga	alabathtta	togotaeaga	acsggcasga	ttaattagty	atgocaatea	960
gs						962

<210> 64 <211> 238 <212> PRT <213> 8.Sapions

<400> 64

Arm Glu Lys The Asp Gln Pro Ser Gly Met Met Pro Phe Cys Nis Asm 1 $^{-10}$

The the Aan the Ser Cys Val Lys Asn Asn Trp Ser Asn Asp Val Arg $20\,$ $30\,$

Ala Ser Leo Tyr Ser Leo Met. Val Leo Hie Lie Leo Thr Thr Leo Val

G)y Asn Leo Ile Val Ile Val Ser Ile Ser His the Lys Glo Leo His 50 60

The Pro The Ash Trp Leu Ile His Ser Met Ala The Val Asp Phe Leu

Let Gly Cys Lou Val Met Pro Tyr Ser Met Val Arg Ser Als Glu His 95 - 95

Cys Trp Tyr Phe Gly Glo Val Phe Cys Lys lle Bis Thr Ser Thr Asp 105

lle Met Leu Ser Ser Ala Ser Ile Phe His Leu Ser Phe Ile Ser Ile 115 Asp Arg Tyr Tyr Ala Val Cys Asp Pro Leu Arg Tyr Lys Ala Lys Met Ash He Leu Val He Cys Val Met He Phe He Ser Trp Ser Val Gro 145 150 160 Ala Val Phe Ala Phe Gly Met Ile Phe Leo Glu Leo Asn Phe Lys Gly Als Gllpha Gllpha Ile Tyr Tyr Lys His Val His Cys Arg Gl γ Gl γ Cys Ser Val Phe Phe Ser Lys Ile Ser Gly Val Leu Thr Phe Met Thr Ser Phe Tyr Ils Pro Gly Ser Ile Wet Leu Cys Val Tyr Tyr Arg Ilo Tyr Leu The Ala Lys Glu Gln Ala Arq Leu The Ser Asp Als Ash Glo <210> 65 <211> 1018 <212> DNA <213> B.Sapiens <400> ascaptopoly quiggsacot gygoatgtat attitigatty tititatycat sciociagig 60 aagaaccast gictigotos gatagaagoa agatactoag actiagitto teigiagoto 120 etgettttta ttattootgg tiggattgoa obactactca gittotatti tataatactg 180 attataaaac atgggaggga aatmactttg tottggtttt totggataat ttattatgtg 240 toctagacto tegeocttyte adaugangya cythagangy cacquistat tatacttegg 300 aatqotagab qaqadkqadd tiggtettiicd andoqqaaqa qqqaaaqqat liitaadlada 360 aatamaqqaa tooagcaqat qqcatoaqaq sacadtataa saaaqaasoq atttqcaada 420 queactions thousasses attentiant totatagets goasgoogt tittigsate 480 540 qaacaqsaca taqtaatata qqaasacaca atqutqagaa aagccaqcaa gttcacacct gtiggggaas ageacacttt teecatetea ggcqtaaeaa toaacagbaa eablacbutg 600 gtacaggttq aqtatocott accessastg titgassess gassigitti ggatticggs 560 ttinggaata ittacacatt cataatgata tatotiggaa atggticcca agictamaca 720 cassattiat tiatettica tatacacett atacacetey tolgasague attligiaca 780 atattttaaa taattttogg vatgassoas agtitgosta cattgascoa toagacagoa 340 awagettewn qtqtqqaxit iteeactigi ggeateatgi igatgeteaa aaagtteeat 900 Page 39

attttagage atttesaatt tiggattite saattaesaa teettaseet giaettagal 960 gitsaataes gigeetette esegggeset tiesggaage atteittist ataageee 1008

<210> 66 <211> 327 <212> PRT

<213> Н.Sapieлы

<400> 66

Tyr lle Lys Glu Cys Phe Leu Lys Vsl Pro Val Glu Glu Ala Leu Tyr l 5 10 15

Leo Thr Ser Lys Tyr Arg Leo Ser Ile Cys Ash Leo Lys Ile Gln Ash $20 \ 25 \ 30$

Lee bys Cys Scr Lys lie Trp Asn Phe Lee Ser Lie Asn Met Met Pro $35 - 40 - 45 \,$

Gin Val Glu Asn Ser Thr Pro Glu Ala Phe Ala Val Trp Phe Asn Val 50 55 60

Cys Lys Leu Cys Phe Mat Pro Lys Ile Ile Asn Ile Val Gln Asn Tyr 65 70 80

Phe Gln Thr Met Cys lle Arg Cys lle Aso lle Aso Lys Phe Cys Val 85 90 95

Thr Trp Glu Pro Phe Pro Arg Tyr Ile Ile Met Aan Val 11e Phe Arg 100 105 120

Asn Pro Lys Ser Lys Thr Phe Leu Vol Ser Asn The Leu Gly Lys Gly 115 120 125

Tyr Ser Thr Cys Thr Thr Va) (10 Leu Leu Lou Thr Pha Thr Pro Glu

Mot Leu Lys $V_{\rm K}$). Cys Phe Ser Pro Thr Gly Val Asn Leu Leu Ala Phe 145 150 155 160

Leo lie lie Val Phe Ser Tyr Ile Thr Met Phe Cys Ser Ile Gln Lys 165 170 175

Thr Ala Leu Gln Thr Thr Glu Val Arg Asn Cys Phe Gly Arg Clu Val 180 185 190

Ala Val Ala Ash Arg Pho Phe Phe IIc Val Phe Ser Asp Ala IIe Cys 195 200 205

Trp Ile Pro Val Phe Val Val Lys Ile Leu Ser Leu Phe Arg Val Glu 210 220

The Pro Gly Glm Ser Leu Leu Ser Phe Pro Ser Lie File His Arg Ala 225 236 236 240

Phe Leu Arg Pro Sor Phe Asp Lye Ale Arg Val Asp Thr Ile Ile His Page 40

245	250	255	
Lys Asm Gln Tyr Lys Val Ile			
260	265	270	
lle Lys Lys Led Ser Ser Gly 275	Ala Ile Glm Pro 280	Gly lie lie Lys 285	. Sen
Arg Ser Tyr Arg Glu Thr Lys 290 295	Ser Glu Tyr Leo	Ala Ser Ile Ala 300	. Arg
Ris Trp Phe Phe Thr Arg Ser 305	Met His Lys Thr 315	lle Lys lle Tyr	Met 320
Pro Arg Pho Miy Pro Gly Len 325			
<210> 67 <211> 1251 <212> UNA <213> H.Sapiens			
<400) 67 aptaccateg magstyacet wester	cantaques Joses	uccacacaga gott	gatgat 60
pappachock accordaage tegete			
gggetgerag cesstgggtt gatgce			
ggeacgeqte tggagetget untget			
goagoggoot todayatoot arayah			
gsetgeeget totaclacki colatg			
gengeretea grategareg etgeet	getą gegetątące	cacactgota cont	. ရေးမှ နေသူမှုနှင့်
equecagted gootgoocet etgggt	etge geeggtgtet	gggtgetgge cace	etette 480
zgegtűeset gyetőgtett csesga	gget geogtetget	ggtacgacet ggto	atotgo 540
ctggacttot gggacagtça qçaqot	gtog otgaçqatço	tggøgetest gggg	iggette 600
etgeetttee toobholgat agtata	grade gtyrtbacer	aggopacago chgt	ogoson 660
tgocaroquo ಶಕ್ಷಕ್ಕೂತ್ರದ ೧೬೧೩೪೮	oston ದರ್ಭಧಿಕ್ಕಾರಿದ್ದಾ	congrutogo cago	Saccatt 7000
etglesgent elgtgetekt gagget	gede taddagetçç	cocagetget ctac	etygee 780
ttoctytgyg addtetadto tggota	cotg ototoggaçç	coctggteta etco	ngadtad 840
ctgatectae teaacagotg cotoag	poers theetetyes	tralignees too	garote 900
eggacectge tgegoteryb nebeko	qtoo ttngcqqcag	otototgoga ggag	oggoog 960
ggeagettea egereactya gecaca	gaco cagotagatt	otgagggtoc aact	etgeca 1020
ფადელც ა ნეფ დ ავაცე<i>დი</i>და ფხია და	gatg gateetgigg	occanoctoa gata	gaaccee 1080

1200

1251

acactocago caegaliugga teorocagot cagocacago tgaaccetae ggoccagoca caghiggate cearageona eccaragete asceteatyg cocageosea gioagattet giggocomen cacaggongs carimaegic cagacocolo cacobhetgo o <210> 68 417 <211> <212> PRT H.Sapiens <213> <400> The The Met Glu Ala Asp Leu Gly Ala Tor Gly His Atg Pro Arg The Glu Leo Asp Asp Glu Asp Ser Tyr Pro Gln Gly Gly Trp Asp Thr Vol Phe Leu Val Ala Leu Leu Leu Leu Gly Leu Ero Ala Asn Gly Leu Met Ala Trp leu Ala Cly Ser Glm Ala Arg His Gly Ala Gly Thr Arg Leu Ale Leo Leo Leo Ser Leo Ala Leo Ser Asp Phe Leo Phe Loo Ala 65 70 75 60 Ala Ala Ala Phe Gln Ile Leu Glu Ile Arg His Gly Gly His Trp Pro Leu Gly Thr Ala Ala Cys Arg Pho Tyr Tyr Phe Leu Trp Cly Val Ser Tyr Ser Ser Gly Leo Phe Leo Leo Ala Ala Leo Ser Leo Asp Arg Cys Leu Leu Ala Leu Cys Pro His Trp Tyr Pro Gly Bis Arg Pro Val Arg Leu Pro Loo Trp Val Cys Ala Gly Val Trp Val Lou Ala Thr Leu Phe 145 150 150 Sor Val Pro Trp Loc Val Pho Pro Glo Aka Ala Val Trp Trp Tyr Asp Leu Wal lie Cys Leu Asp Phe Trp Asp Ser Glu Glu Lou Sor Lou Arg Met Leu Glu Val Leu Gly Gly Phe Leu Pro Phe Leu Leu Leu Leu Val Cys His Wal Leu Thr Glo Ale Thr Ale Cys Arg Thr Cys His Arg Cln

Gin Gin Pro Ala Ala Cys Arg Gly Phe Ala Arg Val Ala Arg Thr Ite

Leo Ser Ala	Tyr Val Val 245	Leu Arg	Leu Pro 250	Туг	Gin Leu		in Leu 65		
Leu Tyr Leu	Ala Phe Leu 260	Trp Asp	Val Tyr 265	Ser (Gly Tyr	Leu La 270	en Trp		
Glu Ala Leu 275	Val Tyr Sei	Asp Tyr 200	Leu Ile	Leu :	Leu Asn 285	Ser C	ya Leu		
Ser Pro Pho 290	Lou Cys Leu	Met Ala 295	Ser Ala		Len Arg 300	Thr L	eu Leu		
Arg Ser Val 305	Leu Ser Ser 310		Ala Ala	њес (315	Cys Glu	Glu A	rg Pro 320		
Gly Ser Che	Thi Pro Thi 325	Glu Pro	Gln Thr 330	Gln I	Leu Asp		lu Gly 35		
Pro Thr Leu	Pro Clo Pro 340	Met Ala	Glu Al a 3 4 5	Gln !		Met A: 350	sp Pre		
Val Ala Glo 355	Pro Gla Val	Asa Pro 360	The Leo	Gin I	Pro Arg 365	Ser A:	ap Pro		
Thr Ala Glo 370	Pro Gln Lev	Asm Pro 375	Thr Ala	Gln i	Fro Gln 380	Ser A	SP Pro		
Thr Ala Glo 305	Pro Gln Leu 390		Met Ala	Gln 1 395	Pro Gln	Ser A	ep Ser 400		
Val Ala Gln	Pro GLo Als 405	Asp The	Asn Val 410	Gln !	Thr Pro		ro Ala 15		
Ala									
<210> 69 <211> 659 <212> DNA <213> U.Sap	<211> 659 <212> DNA								
<400> 69 tacaggedtg a	geatgetgg g	ctccateag	; caccasç	jeac 1	tgaaligta	ca te	støtggee 60		
catetagtae e	getgecape a	ccccacaca	octgics	igea (gtegtata	to obs	yatatgeg 320		
ecatgiacet g	otgoagago a	teetggaat	: ggatgtt	etg 1	tggettee	tg te	agtggtg 180		
etgattetgt t									
gligggillat a	tigogggbee s	gedegijt te	: Egetggt	icag i	رة المصاهي	gt eg	stocegga 300		
agatgoddt g	accaggotg t	acatgacca	tootgat	cag .	agtgctgg	te tt	retectot 360		
gtgacctgcc c	ettiggdati d	agtgattco	: tatttt	ctg ·	gatocacg	tg ga	tttgtcac 420		
gttogtotag t	trocattit o	etgtepact	: ottaaca	igoa i	gtgccaac	co ca	ttatttac 480		
ttetteatgg g	otectting q	cașetteac		ාසයුත : බලුළ 4		ंद वृत्	ttotocag 540		

600 659

agggetetme aggacacqee iqaggtqqaa qaagqcaqat ggeggettte tqaqqaaacc etggagotyt detgsageag attggggdda tgaggsageg cololgcdot gtgabboeg <210> <211> 213 <212> PRT <213> H.Sapiens <400> 70 Tyr Arg Pro Glu His Ala Gly Leu His Gln His Gln Ala Leu Fro Val His Pro Val Ala His Leu Val Pro Leu Pro Pro Pro His Thr Pro Val 20 25 30 Ser Ser Arg Val Ser Cys Ser Cly Pro Cys Pro Cys Cys Arg Als Ser 35 40 45 Trp Asn Cly Cys Sar Va) Ala Ser Cys Loo Val Val Leu lle Leu Phe 50 55 60 Gly Vel Lys His Gln lle Ser Ser Gln Ser Hia Gly Phe Phe Tyr Val 65 70 80 Trp The Ser Ala Gly Pro Ala Arg Phe Cys Trp Ser Gly Ser Phe Val Asp Pro Cly Ard Cys Pro Pro Cly Cys Thr Pro Ser Cys Ser Glu Cys Trp Ser Ser Ser Val Thr Cys Fro Leu Ala Phe Ser Asp Ser Tyr 115 120 125 Phe Ser Gly Ser Thr Trp Ile Cys His Val Arg Leu Val Ser Ile Phe Lou Son Thr Lou Asa Sen Ser Ala Asa Pro Ilo Ile Tyr Phe Phe Met Gly Ser Fhe Arg Gln Leu Gln Asn Arg Lys Thr Leu Leo Vai Leo Gln Arg Ala Leu Gln Asp Thr Cro Glu Val Glu Glu Gly Arg Trp Arg Leu Ser Clu Clu Thr Leu Glu Leu Sor Sor Arg Leu Gly Pro Gly Arg Ala Ser Als Leu Ser Val $\mathbb{R} \mathbf{10}$ <2105 71 5.59 <211> <232> DNA

Page: 44

<213> H.Sapiens

<400> 71 atgccgaagg	caggeegeag	១១៤ ១៤៤១១៤៤៦	gaggaeggtg	acgaggatga	феревара д аа	60
gecoeggegt	ggggccgct	ggggaatig	otocanc o ga	agcageagea	təaggatgga	120
cocacacati	gtgcaacaca	gnagagedag	dagoeddgot.	gooscoagoc	acadedteed	180
gcadaagtgg	eggetggget	ccccgaagaa	ctgggtgcag	qegoogetga	gcagcagistg	240
pagpagpagg	cagagggccc	aggtgaggge	gbacababag	gtggtcaggt	ggagtgggag	300
debdevečeć	taccaggotg	qqaaqauçqe	9900иврово	tgeteeaege	tgaoggoogo	360
caggaçanto	ауроспасца	tgtageaças	gaagogcago	gtfgodaged	tegtotqoad	420
gaagecoggg	aayteeagee	ggmottycag	caagtegggg	acqatqqccs	ccatqtqyca	480
gecaaggaag	atgagatoog	cqcaggccac	gtocaggagg	tagatggcça	aagggtblot	540
gtagadattg	gagotgago					559

<21/0> 72 <211> 213

<21:> PRT <213> N.Sapiens

<400> 72

Leu Ser Ser Asn Val Tyr Arg Asn Pro Phe Als Ile Tyr Leu Leu Asp

Val Ala Cya Ala Asp Leo Ile Phe Leo Cly Cya His Met Val Ala Ile 20 25 30

Vai Pro Asp Leu Leu Gin Gly Arg Leu Asp Phe Pro Gly Phe Val Cla

The Ser Leu Ala Thr Lau Arg The Che Cys Tyr Tie Val Gly Leu Ser 55 50

Lou Let Ala Ala Vat Sor Val Glo Gin Cyc Lou Ala Ala Let Fhe Pro

Als Tip Tyr Ser Cys Arg Arg Fro Arg His Len Thr Thr Cys Vel Cys 85 - 90

Als ben Thr Trp Ala Leu Cys ben Leu Leu His Leu Thr Thr Cys Val

Cya Ala Leu Thr Trp Alo Leu Cys Leu Leu Leu His Leu Leu Leu Sor

Gly Ala Cys Thr Leo Leo Leo Ser Gly Ala Cys Thr Glo Phe Phe Gly

Glu Pro Ser Arg His Lew Cys Arg Thr Lew Trp Lew Val Ala Ala Val

Leu Leu Ala Leu Cys Cys Thr Mot Cys Gly Ala Ser Leu Met Leu 165 170 175

Leu Leu Arg Val Glu Arg Gly Fro Gln Arg Pro Pro Pro Arg Gly Phe 180 185 190

Pro Gly Leu Ile Leu Leu Thr Val Leu Leu Phe Ser Ser Ala Ala Cys 195 200 205

Leo Arg His 210

<210> 73

<211> 1008 <212> DNA

<213 × B.Sapiens

<400> 73

atggamtest efficient typagtrute effectetee tygesteest catestryst 60 actawcacwo tagtagoolgi qqohqteeba otqibyatoo acaegaatga tggtgtcagt 120 cletgetica cettgaatet ggetgtgeek gadaddibga blogtgtgge catototgge 180 ctactoscag accagetete cagecettet eggeocacae ageagaceut ytgesgeetg 240 oggatggdat tigicacito otcogoagei geetetgice teaeggicat geigaldade 300 tttgadaggt additgddat daagdagddd ttddgdtact igaagaidai gagiqyg. Co 360 gtiggeoggeg estgemitige egypetigigg thagigiett accidating etteologia 920 ologgaahoo qoalgChoca gobÿaclgoo lacaaayeyo agtgoagott offtgotgta 480 titescecte acticgiget gardeteted igegiigget keiltocompo catquitecte 540 titqtottot totactgoga catgotoaag attgootoca tgcacagora geagabtoga 600 aagatggaac atgeaggago catggotgga ggttatogat coocaeggac teecagcçac 660 tionmagete teograpist giologistote attgggaget tigetetate eiggaeecee 720 clocitalea etgycatigi gewyglygee tyecaegagt yteaceteta cetagigety 780 gaseggtace tytygotyct egycytyyyc aactroolyc teamcecach calotatroo 840 tattggcaga aggaggtgog actgcageto taccacatgg coctaggagt gsagaagglg 900 960 etenectrat tectectott tetotoggod aggaattgtg goodagagag goodagggaa 1008 amblections acategorese taletecage heagagilite atogetea

<21.0> 74

<21.1> 335

<212> PRT

^{₹213&}gt; H.Sapiena

<400> 74

Met Clu Ser Ser Pho Ser Pho Cly Val IIe Leo Ala Val Leo Ala Ser 1 5 10 15 Let I.ic IIe Als Thr Aso The Let Vol Ala Val Ala Val Let Let 20 25 30The His Lys Asn Asp Gly Val Ser Leu Cym Phe Thr Leu Asn Leu Ala 35 40 45Val Ala Asp Thr Leu He Gly Val Ala He Ser Gly Leu Leu Thr Asp 50 60 Glo Leu Sor Ser Pro Ser Arg Pro Thr Gln Lys Thr Leu Cys Ser Leu 45 70 75 80 Arg Met Ala Phe Val Thr Ser Ser Ala Ala Ala Ser Val Leu Thr Val Met Lau Ile Thi Phe Asp Arg Tyr Leu Ala lle Lys Glm Pro Phe Arg Tyr Lea Lys 11ê Met Ser Gly Pho Val Ala Cly Ala Cys Ile Ala Gly 115 120 125 Len Trp Leo Val Ser Tyr Len Ile Gly Fhe Lou Pro Leu Gly Ile Pro 135 Met Phe Glm Glm Thr Ala Tyr Lys Gly Glm Cys Ser Phe Pho Ala Val Phe His Pro His Phe Val Leu Thr Leu Ser Cys Val Gly Phe Phe Pro Als Mot led Leo Phe Val Pho Pho Tyr Cys Asp Met Leo Lys Ile Ala 180 185 190 Ser Met His Ser Gln Gln Ile Arg Lys Met Glo His Ala Gly Ala Met Ala Gly Gly Tyr Arg Ser Pro Arg Thr Pro Ser Asp Phe Lys Ala Leu Acq Thr Val Sec Val Lan Ile Gly Ser Pho Ala Lau Ser Trp Thr Pro The Leo fle Thr Gly Tle Val Glo Val Ala Cya Glo Glo Cya His Lou 245 250 255Tyr Leu Val Leu Glu Arg Tyr Leu Trp Leu Leu Gly Val Gly Asn Ser 265 270 Leo Leo Asn Pro Lou Ile Tyr Ala Tyr Trp Cln Lys Clu Val Arg Leo Gin Leo Tyr His Met Ala Leo Gly Val Lys Lys Val Leo Thr Ser Pho 290 295 300

Leu leu Phe leu Ser Ala Arg Aso Cya Gly Pro Glu Arg Pro Arg Glu 305 --- 310 --- 315 --- 320

Ser Ser Cym His Ile Val Thr Ile Ser Ser Ser Glu Pho Asp Gly 325 330 330

<210> 75 <211> 2137 <212> ENA

<013> H.Sapions

<4002 75 60 aactqqaaqq qeageegtet geegeecacq aacacebtet chageacttt gagtgaccae gmottgcaag etggtggotg geocoocgag tecogggoto tgungomogg cogtogastt 120 160 apportigea tectgitace tggagacect etgagetete andigetact tetgeogotg offichiqueen gageocoqqqe gaggacocot coaggatgca ggtcccgase aqescogque 240 eggacaange gacgetinae algetinega accegnogat egengtiges etileseigtige 360 360 tgtactogot ggtggoggog gtcagoatoc ogggomanch utbutetetg tgggtgotgt googgogoat ggggoodaga toooogtogg teatetteat gatqascobq aquettacqq 42°D acctgatget egocagogty tigoettico aastetacta cestigosad egucachaet 489 540 englablicen represents the another tracestage cittlegges ascattiall edagoalock cachalyson lykatesges typassett cotsgesste otstacees: E(0)teageteeas gegetggege egeogtegtt accogginge vgegintgia gggaeeigge 660 720 tgotgotoot gacogodotg teocogotgg cycgesecga Leternolae coggigoacg cootoggical cateacetge ttogacytes teaaytiggae qatgateeco aqaqliqqeea **780** tetagecont attactatte accatettes testgetatt accesteery itogipates 840 complayed to the the angles assats of the age to the change of the control of the 900 gygaydagog gaggogogog ghgggoddog cogogotggt ettgetggod titgicaeet 960 1020 gettegecec caacaactte gigeteeigg egnacabegt gageogooig tictaeggea 1080 agagetacta coacqtgtac aageteacge tgtgteteag etgeeleaan Aantqtetge accognible thattauthi gogtocopgy authoraget gegeotgogg gaatattigg 1140 gatgaagaay ggtgaadaya yacacaatgg acaagagaan agaaagaata ttotoogooa 1200 1260 gyaccaegte egtgegetee gaggeeggtg egeaccetga agggatggag gyagbuacua ggeceggest ceagaggeag gagagtgtgt tetgagtese gggggegeag ettggagags 1320 ogggggggga gettggagga tebagggggg catggagagg ccaeggtgec agaggttcag 13B0 1440 ggagaacage tgoqttgeto compgeacty covangeess gtggggaagg qtotocagge Page 46

tttatteete	coaggraptg	cagaggcacc	ggtgaggaag	ggtotocsgg	cttcactcag	1500
ggtagagaaa	саадсанадс	costeadada	acanggitget	tgttatoctg	cagagggtgc	1560
utolgostot	etgtgt:agg	gescagetty	tqCoaccacg	acaggotaat	ttttgtattt	1620
thithertag	agetgggetg	\$cacccccga	ectecttaga	caetosteas	acototocat	1680
secegaçgat	ggatattcaa	ccagcoccac	ogcotaccny	satinggibbe	tggalatoot	1740
etgtgggega	actgogagod	ccattoccag	ctattatass	tgctgacatc	gtocctteec	1000
acacctytec	atacccgagg	atggatatte	aaccageccc	accycctacc	cgactoggtt	1,860
totogatate	ctatgtegge	qaactgogag	coccattocc	agetettete	cctgotgaca	1920
hogtpoolits	gtEgEgette	Eggeotte()	callectoctc	caqqqqttct	ggtotoogta	1960
geeoggtgea	egeogaaatt	totgtttatt	toactcaggg	gcactgtggt	tgatqlagtt	2040
ggsattette	tttcagagga	gagaat ggag	ctootgoaag	teagetacte	toogtgooca	2100
etteesetsa	Офествовор	cocotogtgo	ogeatte			2137

<210> 76 <211> 359 <212> PRT

<400> 76

Met Gln Val Bro Asn Ser Thr Gly Pro Asp Asm Ala Thr Leu Gln Met 1 5 10

Leu Arg Ash Pro Ala IJe Ala Val Ala Leu Pro Val Val Tyr Ser Leu 20 25 30

Val Ala Ala Val Ser Ile Pro Gly Asn Leu Phe Sec Lou Trp Val Lau

Cys Arg Arg Met Gly Pro Arg Ser Pro Ser Val Ile Ehe Met Ile Asn 50 55 60

led Ser Val Thr Asp Led Met Led Ala Ser Val Led Pro Phe Glm 11e 65 75 80

Tyr Tyr His Cys Asn Arg His His Trp Val Phe Gly Val Lou Lou Cys

Ash val Val Thr Val Ala She Tyr Ala Ash Met Tyr Ser Ser Ile Leb

Thr Met Thr Cys Ile Sor Val Glo Arg Phs Lou Gly Val Leu Tyr Pro

Leo Sor Ser Lys Arg Trp Arg Arg Arg Arg Tyr Ala Val Ala Ala Cys

fage 49

<238> N.Sapiene

Ala 145	Gly	The	Тгр	Len	Leu 150	Leu	Leu	Thr	Ala	Leu 155	Ser	Pro	Len	Ala	Acg 160	
Thr	qeA	Leu	Thr	Tyr 165	Pro	Va.1	His	Ala	100 170	Gly	11e	Ile	Thr	Су <i>в</i> 175	Fhe	
ħsp	Val	Leu	Lys 180	Դ. Հեն	Thr	Met	Leu	Pro 385	Ser	Val	Ala	Met	Trp 190	Ala	Val	
Fhe	Leu	Phe 195	Thr	Ile	Phe	11 e	Leu 200	Leu	Fhe	T _{elect} a	11¢	Pro 205	Ph⊚	Vo.1	Ije	
Thr	Val 210	Ala	Сув	Tyr	Thr	Ala 215	Thr	11e	Leu	Lys	Leu 220	Leu	Ang	Thr	G10	
61u 325	Ala	Ris	Gly	Arg	Gl u 230	G1n	Arg	Arg	Arg	Ala 235	Val	G1 y	Ţēņ	Ala	Ala 240	
Vel	LSV	Les	Leu	Ale 245	Phe	Vel	Thr	Cys	Phe 250	Ala	Рго	āso	Asn	Phe 255	Val	
ьео	Leu	Ala	His 260	lle	Val	Ser	Arg	Leu 265	Phe	туг	Gly	ьуг	Ser 270	Tyr	Tyr	
Нав	Val	Tyr 275	Lys	ren	Thr	Leu	Cys 280	Leu	Ser	Сув	Leu	Aan 285	Aan	Сув	Leu	
Asp	Pro 290	Phe	vel	Tyr	Tyr	Phe 295	Als	Ser	Arg	GJ u	Phe 300	Gin	Leu	₽±ij	Leu	
Arg 305	Glu	Туг	Leu	Gly	Сув 310	Aug	Arg	Val	Pro	Arg 315	Агр	The	Leu	Asp	The 320	
Arg	Arg	Glu	Ser	Նөս 325	Phe	Ser	Ala	Arg	Thr 330	Thr	Ser	Val	Arg	Ser 335	Glu	
Ala	СТУ	Ala	Bis 340	Pro	Clu	Gly	Met	G1u 345	Gly	Ala	Thr	Arg	Pro 350	üly	Leu	
Gln	ħrġ	Gl.n 355	GLu	Ser	Val	Phe										
<21 <21 <21 <21	1> 2>	77 1197 DNA B. So	pien	5												
<40	0> oant	73 cno	oact.	catio	ം വ	იბაი	04100	e at	ರಾಗರ	ខេត្តព	tea	togt	set	geat	tacesc	60
															gaegee	120
															ittggLg	100
															aogttg	240
															introca.	300
t e é	ស្រួស គេ	Legis	rüğt	agg.	⊸ ⊃و	. ug te '- 1		4.						•	- **	

PCT/US00/31581 WO 01/36473

otgaaactgt coccegeget	etgątteęca	edàāadāāaā	gegtettegt	ggcaetcact	360
genteegine igaqueteet	ggccategeg	et geagogea	gootbaccat	ggogogoagg	d 20
ggyddegena begtetebad	ე იტმტმმიბი	<i>ងប់</i> មុខ៦ផ្លួចចូ <i>ង</i>	tggcagcogc	ggootgaga	480
glytegatys tostogggal	popeoopios	atqqqotqqa	attgootggq	togootogaa	540
gettgetons etgtettged	getetaegee	asggootadg	tigatet teta	cqtgotogco	600
ttogtgggda teetggeogd	tatotgtgca	ctotacgogo	geatotactq	chaggiacge	660
decendade adeacatace	ååcecååcec	gggactgcgg	ggaccacctc	gsaccqggaag	720
egtegraage egegeteget	ggoottgatg	ogoaogotica	gegtgatgat	octggodttt	780
etggoutgit geegoooct	attoetgatg	etatinetog	acqtogogtg	popagagagage	840
acciginate tactoriges	ggoogatoco	Rtectgggen	tggoostggo	caactcactt	900
ctgascccea tratetacac	getcaccaac	o go ga dotaje	gedacgegut.	adligageetg	960
gtotgotgog gaogecacto	otgoggoaga	gaccogagtg	gotoccagoa	gteggegage	1020
geagetgaag etteegaaga	entgageege	tgaatgaaco	egggeettga	tyggagette	1080
ageggetegg agegetestu	çconcagogo	gaoggqotgq	поассадеру	otopadaggo	1140
agcoccygty caccescage	egeceggaet	otggtatcag	ascogeobge	agaotga	1197

<210> 78 <211> 398 <212> PRT <213> B. Sapiens

<400> 7~

Met Gin Ger Gly Leo Leo Arg Pro Ale Pro Val. Ser Glo Val. He Val. 1 5 10 15

Leo His Tyr Ash Tyr Thr Gly Lys Leo Arg Gly Ala Arg Tyr Glo Pro $20 \ \ 20$

Gly Ala Gly Leo Arg Ala Asp Ala Val Val Cys Leo Ala Val Cys Ala

Phe Ile Wat Leu Glu han Lou Ala Val Leu Lau Val Leu Gly Arg His

Pro Arg Phe Bia Ala Pro Met Phe Leu Leu Leu Giy Ser Leu Thr Lou 65 76 80

Ser Asp Leu Leu Ala Gly Ala Ala Tyr Ala Ala Asn Ile Leu Leu Ser 85 90 95

Cly Pro Leu Thr Leu Lys Leu Sur Pro Ala Leu Trp Pho Ala Arg Glu

Gly Gly Val Phe Val Als Leu Thr Ala Ser Val Leu Ser Leu Leu Ale lle Ala Leo Glo Arg Ser Leo Thr Met Ala Arg Arg Gly Pro Ala Pro Val Ser Ser Arg Cly Arg The Leu Ala Met Ala Ala Ala Ala Trp Gly Val Ser Lou Leu Leu Gly Leu Leu Pro Ala Leu Gly Trp Asn Cys Leu Gly Arg Len Asp Ala Cys Ser Thr Val Len Pro Len Tyr Ala Lys Ala Tyr Val Leu Phe Cys Val Leo Ala Phe Val Gly Ile Leo Als Als Ile Cys Ale Leu Tyr Ala Arc Ile Tyr Cys Cln Val Arg Ale Aso Ale Arg Arg Leu Pro Ala Arg Pro Gly Thr Ala Gly Thr The Sec The Arg Ala Arg Arg Lys Pro Arg Ser Leu Ala Leu Leu Arg Thr Leu Ser Val Val Leu Leu Ala Phe Val Ala Cys Trp Gly Pro Leu Phe Leu Leu Leu 200 265 270 Leu Asp Vet Ala Cys Pro Ala Arg Thr Cys Pro Val Leu Leu Gin Ala Asp Pro Phe Leu Gly Leu Ala Met Ala Asn Sar Leu Leu Aan Pro Ile lle Tyr Thr Leu Thr Ash Arg Asp Leu Arg Ris Ale Leu Leu Arg Leu Wal Cys Cys Gly And His Ser Cys Gly Ang Asp Pro Ser Gly Ser Gln Gio Set Ala Ser Ala Ala Glu Ala Ser Gly Gly Leu Arg Arg Cys Leu Fro Fro Gly Leu Asp Gly Ser Dhe Ser Gly Ser Glu Arg Ser Ser Pro Glm Arg Asp Gly Leu Asp Thr Ser Cly Ser Thr Gly Ser Pro Gly Ala Pro Thr Ala Ala Arg Thr Leo Val Ser Glu Pro Ala Ala Asp 7.9 <210>

Page 52

<211>

<212>

1041

DNA <213> H.Sapiens

<400> 79						
	ggtcgtgctg	cogcategag	åååäs es cen	tetecesgyt	detacoûncd	60
ctqctcattq	tggcetttgt	getäägeges	otaggoaatg	gggtegeect	gtgtggtt()e	120
tgottopada	tgaagacetg	фавроссвос	actittacc	ttttcaattt	ggaagtigat	780
gatiticotco	ttatgatotg	cetgectitt	oggacagact	attacetcag	acgtagacac	240
tgggettttg	gggacattee	etquoqagte	gagetettea	cgttggccat	gaacagggcc	300
gggagcatog	tgbtccttac	ggtggtggct.	geggacaggt	atttoaaagt	ggtocacocc	360
caccacgegg	tgaacactat	ctocaccegg	gtggeggetg	geategleig	caccolytyg	420
gcootggtca	tootgggsac	agtgtatett	ttüctggaga	accatctctg	ogtgcaagag	4 B.CI
acggoogtot	ootgtgagag	etteateate	gagtoggoda	atggotggoa	tgacateatg	5.4(1
ttocaepatgg	acttetttat	geocatogya	stostettat	tttgeteett	casqattgtt	6 00
tygogratga	ព្រំប្នូលព្រះបាន	geagetages	ogacaggoto	ggatgaagaa	ggogacongg	660
tituat daligg	tiggtiggeast	tgtgtteatd	adatigohade	tycocagoyt	gtotgotage	720
ctctatttcc	tetggaeggt	gecetegagt	gadt gagat a	calchetacea	Legggeeetç	780
cacatasocc	tcagottcap	otacatgasc	ageatgetgg	atococtggt	gtattatti	840
toaagcooct	cettteecaa	attotacaso	payeteapaa	totgoagtot	рвавоосавд	900
(асысөфдас	actoesaesc	асавадосод	gaagegat.gc	caabbhogae	ectoggtage	960
aggaqttgca	tesytytyge	asatagttto	casagocagt	ctgatgggca	atgggatece	1020
racatigity	agtqqcactq	ā				1041

<310: 80

<4005 80

val Met Pro Pro Leu Leu Ile Val Ala Phe Val Leu Gly Ala Leu Gly 20 25 30

Ash Gly Val Ala Lee Cys Gly Phe Cys Phe His Mot Lys Thr Trp Lys $\frac{35}{45}$

 $P_{T^{\prime}J}$ Ser Thr Wal Tyr Leu Phe Ash Leu Als Val Als Asp Phe Leu Leu

Met lle Cya Leu Pro Phe Arg Thr App Tyr Tyr Leu Arg Arg Avg His 65 70 80

<211: 346 <212: PRT <213: U.Sepiens

Trp Ala Phe Gly Asp Ile Pro Cys Arg Val Gly Leu Phe Thr Leu Ala 85 95

Met Asn Arg Als Gly Ser He Val Phe Loo Thr Val Val Als Als Asp 100 105 110

Arg Tyr Phe Lys Val Val His Pro His Die Ale Val Ash The Ite Scr 115 120 125

Thr Arg Val Ala Ala Gly The Val Cys Thr Leu Trp Ala Leu Val The 130 140

Leu Gly Thr Val Tyr Leo Leu Leu Glo Ash Ris Leu Cys Val Gln Glu 145 150 156

Thr Ala Val Ser Cys Glo Ser Pho Tle Met Glu Ser Ala Aso Gly Trp 165 170 175

His Asp Tie Met Phe Gln Len Gln Phe Phe Met Pro Len Gly Ite Ile 180 185 190

Len Phe Cya Ser Phe Lys lle Val Trp Ser Leu Arg Arg Glo Glo 195 200 205

Leu Ala Arg Gln Ala Arg Met Lys Lys Ala Thz Arg Phe Ile Met Val 210 215 220

Vel Ald The Val Phe lie Thr Cys Tyr Leu Pro Ser Val Sec Als Avg 205 230 235 240

Leu Tyr Fhe Leu Trp. Thr Val Pro Ser Ser Als Cys Asp Pro Ser Val 245 255

His Sly Als Leo His Ile Thr Leo Ser Phe Thr Tyr Met Ash Ser Met 200 265 270

Leu Acp Pro Leu Val Tyr Tyr Phe Ser Ser Pro Ser Phe Pro Lys Phe 285

Tyr Asn Lys Leo Lys Ile Cys Sar Leo Lys Pro Lys Gln Pro Gly His 290 295 300

Ser Lys Thr Gln Arg Pro Glu Glu Met Pro Ile Ser Asn Lou Gly Arg 305 310 315

Arg Ser Cys Ile Ser Val Ala Asn Ser Pho Glu Ser Glu Ser Asp Gly 335

Cln Trp Map Pro His lie Val Glu Trp His 340

-210> 81

<211× 2525

4212> DNA

<213> H.Sapiens

 $<\!400\!>\!-81$ caagsatgab aggtgactto ocaagtatge etggeescas taceteesgg satteetett. Page 54

genatectat	agtgacaccc	carttaatca	geetetaett	catagtgott	attggeggge	120
tggtgggtgb	catttccatt	cttttcctcc	tgglqaaaat	gascacecgg	teagtgacca	180
ecatggeggt	cattaacttg	gtygtggtad	acagogitti	totgotyaca	gtgccetttc	240
gottgaosta	cotcatcaag	aagacttgga	tgtttggget	gedettetge	essitt.glige	300
gtgecatget	geacatesae	atqtacetca	cqttootatt	ctatgtggtg	atcotggtcs	360
chagatacot	natoltohto	aaybgcaaag	acaaagtuga	attotacaga	aasotgoatg	420
ctytagetys	cagtgctggc	atytygacyc	tggtgattgk	cattglyggta	occotegttg	480
teteeeggta	tggaatocat	gaggaatacs	atgaggagca	ctgtttteaa	Uttoaceaag	540
agottgotta	cacatatgtç	aaaatcatca	actatatgat	agtcatttt	gtostagopo	KOU
ttgetgtgat	totattagto	ttccaggtot	tcatcattat	gttgatggtg	cagaagetac	66.0
geometettt	actatoccae	caggagttet	gggotcagot	gassascota	ttttttataç	720
gggtnatoot	tyttigtilo	entique bace	anthetttag	gatotattac	tigasigity	780
tgacgcattc	caatgootgt	явся бся в фф	procalitita	taacgaaato	ttettgagte	840
taacagcaat	tagotgotat	gattigette	tetttgtett	tgggggAag:	cattggttta	900
agcaasagst	aattggctta	tgqaattgtg	tittgtgaag	ttagccacaa	actacegtal	äen
teatatitge	ttootttata	ttgggaataa.	aaətgegtat	aggggaggta	agaatggtat	1020
ttoalbactf	gaticamaacc	atgood Eqat	qtacnesaaa	cassangact	ataaaatgca	1080
agagodotda	tigtagtoct	tatąggatoc	eteccatete	tgagtgatgq	coutecaeag	1140
accagtgttç	tigaatocac	ctqqagttqc	aatsttacat	tattttccag	tacagaat.gt.	1.300
etgtgtggad	cat gasages	acataggttt	taagagtttt	agagttteat	tageteatte	1260
teaglitiecto	bgi‼lgeag≎	atoghebett	aggbbttgga	otgaacboag	scotttagtt	1320
ctttlealdd	cactteect	taggtaagta	auttotgycc	восарорафо	tellasagaes	1300
caaectotoc	ttogotaaco	aggttagatg	toccattcat	ctcatgeoct	gatampasol	1440
gataagggga	gagaatagtt	aaasattttt	ctagggtate	ətsactolog	taggzagtoa	15000
tetgtetaga	aatcaagaga	awaaqaacgt	gtiggeetiset	gttataacaa	gggtttctag	1560
attemtoctm	tgaaaggtog	tttaaggast	tggggateas	etteeteaat	tateaccaat	1620
tocactytty	obocsasset	cetttasaeg	ettactgyac	atatotscat	aatogtgaaa	1.600
etqtaattle	qagactatcc	ctgactsatg	tgctggtagg	cattaasatg	agttoccsag	1740
gquagtgalt	aaaattittt	tetettetgt	tttttgagæg	astitctaga	tgteetggge	1800

Page 55

caceçttaat	tasgattttt	agggggaca	gaasgttata	etgsaatett	tagagetese	1860
ttoogoogtt	aasattatat	atatatatat	ttaaattata	cottaagtto	tggggtacat	1920
gtgcagaatg	tgcaggtttg	ttabataggt	atacacgtgc	catggtggtt	tgeggeacet	1960
gtcaacccat	otacattagg	tatttetset	aatgstetee	otocoptago	0000020000	2040
tggaraggos	ocatt g tŋtg	atgtteccet	ccetgtgtce	atgtgttttc	attgttcasc	2100
teaceattal	aagtoagaa(:	atgoggtgtt	tagttulects	btookgtett	agbttgctga	2160
gaatgatggt	ttceaggtts	saattatata	tttttaaata	aatgaaaact	gtgtttttaa	2220
aagaqgactt	ttgagaagta	tatagaaaaa	coattaattt	agactototo	agattaggtt	2280
gestykagaa	gqttttctga	atathtgaag	aqtogatasa	Laaatqtooc	ccaangcaat	2340
.aasatcataa	tootitaaaa	tataggaasa	ateactesty	ggaactaggo	(teate-bog	2400
ggetgaaata	stotytacas	caaactccca	tgacacatyt	ttacctetgt	ascaescolg	2460
cadatgtacc	cotgaactta	aaataaaatt	taaagtataa	taataaaata	atstggstit	25.20
tettt						2525

<210> 92 <211> 312

<212> PRT

<213> H.Sapiens

<4005 82

Met Thr Gly Asp Phe Pro Ser Met Pro Gly His Asm Thr Ser Arg Asm 1 - 10 15

Ser Ser Cys Asp Pro Ile Val Thr Pro Ris Leu Ile Ser Leu Tyr Phe $20 \\ 0 \\ 25$

The Val Low the Gly Gry Leu Val Gly Val He Ser He Leu Phe Leu 35 40 45

Let Val Val Val His Ser Val Phe Let Let The Val Pro Phe Arg Let 65. 70

The Tyr Lea lie Lys Lys Thr Trp Met Phe Gly Lea Pro Phe Cys Lys 85 90 95

She Val Ser Ala Met Leu Bia He Bia Met Tyr Leo Thr She Leo Phe 100 $$100\,$

Tyr Wal Val Ile Leu Val Thr Arg Tyr Loo Ile Phe Phe Lys Cys Lys 115 120 125

Asp Lys Val Cle Phe Tyr Acc Lys Leo His Ala Val Ala Ala Ser Ala Page 56

:	130					135					340					
Gly 1	Met	Trp	Thr	Leu	vel 150	lle	V≥l	lle	Val	v al 155	Pro	Leu	Val	Val	Ser 160	
Arg '	TYC	61 y	11¢	His 165	Glu	Glu	Tyr	neA	Glu 170	Gla	ніз	Сув	Phe	Lys 175	Phe	
Hia∷ I	Lys	Glu	Leu 180	Ala	Tyr	Thr	Туг	Val 185	Lys	Ile	De	Asn	Tyr 190	Mot	11e	
Val 3	Ile	Phe 195	Val	He	Аха	Val	Ala 200	Val	Il∉	Leu	Leg	Val 205	Phe	61 n	Wal	
Phe :	11e 210	Ile	Net	L∉u	t/e t	val 215	Gln	Lуз	Leu	Arg	Ніа 220	Ser	Leu	Гел	Ser	
His 0 225	Sln	Ğlu	Phe	Trp	Ala 230	C ln	Гел	Lys	Asn	Leu 235	Phe	Phe	Ile	Gly	Val 240	
Il⊕ I	Leu	Val	Сув	Phe 245	Lev	Pro	Туг	Glin	Phe 250	Phe	Arr	Ile	Тут	Туг 255	Lea	
Asn N	Val	Val	ምክድ 2:60	His	Ser	Aeri	Ala	0 ya 265	Asn	Ser	Lys	Val	Ala 270	₽he	Tyr	
Asn C	Glu	11e 275	Phe	Leu	Ser	val	767 188	Ala	Ile	Ser	Cys	Туг 285	Asp	Leu	Leu	
Leu l	Phe 290	val	Phe	Gly	Gly	Ser 295	Нiэ	Trp	Phe	Lys	G1n 300	Lys	Il€	Ile	Cly	
Len 1 308	űτρ	asA	Сув	Val	Leu 310	Cys.	Arq									
<2102 <2112 <2123 <2133	> 1 > 0	3 125 NA .Sap	aiena	1												
<400		3					. Fr+	e oba	d leek	++-	Last			Non sá ence	jacaaa	60
	-		-		-										jetgtg	100
			_												gatee	180
-															accte	240
	-		_												:togtt	30°
	-	-					•	•							itagto	360
															getat	420
	-														iggtas	480
			_													541
MECHO	0 34 0	ra c	yac		- åd) CCCC	I COUL	. act	aa y i	.cod	ooge	. a c co	20 1	, k k-010	gaatt	O 4 14

tgeateageg	tgtoctggat	catgocceto	stgtacagog	gtgetgtett	ctacacaggt	600
	atgggetgga					6€0
	atossaactg					720
	ttotgtatgg					780
gaasstactg	gtagcaagac	agastoated	tcagagagtt	acaaagccag	agtqqccaqg	840
	aageagetaa					900
	geattgatte					960
	tttgchgbbg					1000
	acceatygtt					1080
	cagcasccat					1125

<210> 84

<400> 84

Met Ser Ser Asm Ser Ser Leu Leu Val Ala Val Glo Leu Cya Tyr Ala 1 5 10

Asn Val Aen Gly Ser Cys Val Lys Ile Pro Phe Ser Pro Gly Ser Arg 20 25 30

Val Ile Leu Tyr Ilo Val Phe Gly Phe Gly Ala Val Leu Ala Val Phe 35 40 45

Gly Non Law Leu Wal Met lie Ser lie Leu His Phe Lys Gln Leu His 50 60

Ser Pro Thr Ash Phe Leu Val Ala Ser Leu Ala Cys Ala Asp Phe Leu 65 70 75 80

val Gly Val Thr Val Met Pro Phe Ser Met Val Arg Thr Val Glu Ser 85 90 95

Cys Trp Tyr Phe Gly Arg Ser Phe Cys Thr Phe His Thr Cys Cys Asp 100 105 110

Val Ala Fhe Cya Tyr Ser Ser Leu Phe His Leu Cys Phe Ile Ser Ile 115 120 125

Asp Arg Tyr Ile Ala Wel Thr Asp Pro Leo Vel Tyr Pro Thr Lys Phe 130 135

Thr Val Ser Val Sec Gly The Cys The Ser Val Ser Trp The Leu Pro 145 150 156

led Met Tyr Ser Gly Ala val Phe Tyr Thr Gly Val Tyr Asp Asp Gly Page 58

<211> 345

<212> PBT

<2330 H.Smpiene

				165					170					175	
Trou	Glu	Clu	J.eu 180	Ser	Asp	Ala	Leu	Aen 185	Суя	Пlе	Gly	Gly	0ys 190	Gln	Thr
Val	Val,	Asn 195	៤ 1២	Asn	Ттр	Val	Leu 200	Thr	Asp	Phe	Leu	Ser 205	Phe	Phe	Ile
Pro	Thr 210	Fhe	Ile	мес	Ile	Ile 215	Leu	Туг	Gly	Asn	11e 220	Phe	Leu	Val	Ala
Arg 225	Arg	Gln	Alæ	Lys	Ly5 230	Tão	G1u	रोड़ा	Thr	Gly 235	Ser	Lys	The	Glø	Sor 240
Ş¢:r	Ser	Glu	Ser	Tyr 245	ьуз	Ala	Arg	Val	Ala 250	Arg	Arg	Glu	Arg	Lys 255	Ala
Als	ьуя	Thr	Leu 26 0	Gly	Val	Thr	Val	Val 265	Ala	Phe	Met	Ile	Ser 270	Trp.	Leu
Pro	Tyr	Ser 275	Ile	Asp	Ser	Leo	580 11e	Asp	Ala	P)tra	Mot	G) y 205	Phe	Ilc.	Thr
Pro	Ala 290	Сув	Ile	туг	Glu	11e 295	Сув	Сув	Trp	Сув	Ala 300	Tyr	Tyr	Asn	Ser
ala 305	Met	Asn	Pro	Leu	lle 310	Туг	Ala	Leu	Phe	Туг 315	Рго	Trp	Phe	Arç	Lys 320
ala -	Ile	Ĺув	Val	Ile 325	Val	Thr	Gly	Gln	Val 330	Leu	Lys	Aso	Ser	Ser 335	Ala
The	Mot	Азл	Lee 340	₽he	Ser	G10	His	11e 345							
<210 <210 <210 <210	1> : 2> (85 1020 MA 4. Sap	oiens	3											
<400 a c c s		85 atgla	igoda	ectaç	a ot	attt	agos	ı aat	gett	etg	attt	ouce	ega t	tatç:	poaget
gett	ttg	jaa a	ittņo	eacts	ea ta	;00a;	ecator	2.004	ietea	ទេជិច	tges	etac	sot t	eo et g	Stbatt
tate	geat	itë t	ot to	enthog	t. gg	gett	teca	a ggo	raatç	jcag	tagt	gata	ste o	caett	iacatt
Litina	aaaat	်ရွာချင့်	gscot	tgga	ടെ മൂട	gcag	jese:	e ato	atta	itge	tgaa	icet	go o	stgca	ıcagat
etge	etgta	atc t	gado	sage:	et ed	ectt	cety	, att	caet	act	atgo	ee a git	्वातु ५	ខ្លួតគត	acteg

atottiggag atttoatgig taagblikabo ognilnagel toratticaa ootgiabago

accatected lockcacety titrageate throughtact gigigatest tracecasts

agetge(tt); ceattesess asetegatgt geagttgtag cetgtgetgt ggtgtggate

alliteactgg tagetgtest teegatgace thettgatem catesacesa esegacesac

Page 59

360

420 480

54Û

gletegeest	caccagtteg	gatgaactca	atactattaa	gtggtacaac	600
ctyceaglac	tttetgeete	ccettggtga	tagtgacact	ttgctatacc	660
acacteegad	ccatogactq	сарастраса	gotgoottaa	goagaaagca	720
coattolyet	actocttyca	tttttacytab	gttttttacc	ettecatate	780
ttdaggatog	aatotoagoo	tgettteast	cagLtgttcc	attgagaatd	840
agettacate	gtttctagac	cattatgctg	ctctgaacsc	ctttggtaac	900
atgtggtggt	саўсцараас	tttcagcagg	etgtetgete	aacagtgaga	960
					1020
	otycsaglad acacttigad coattolgab troaggateg agettacate atgtogtggt	ctycsaglae thletgeole acactifyae coafegaetg coattolgel actentigea tteaggateg aateteagee agettacate gittetagae atgiggigg cagogaeaae	otycsaglae thlotycobe ecottyptga acaettigae ecategaety caanetgaea coattolgeb actentigea thibacytab thraggateg aatereagee tyetteaat agettacate gittetagae cattatgety atgiggtyy cagogacaae threageag	ctycsaglae titetgeete ecettgetga tagigacaet acaettegae ecaregaetg caaactgaca getgeettaa ceattetget actentigea tillacqtai gittittace ticaggateg aatereagee tgetticaat eagligitee agettacate gittetagae cattaigetg etetgaacse atgiggigg cagagacaae titeagpagg etgiotgete	quetoquect excessites gateanetes statitus singularease etgesagiae titeracete ecetigens taginacae titerature scaetiums ecanomics acteriums tuttacetae necetigens cattaline necetigens cattaline necetigens cattaline necetigens cattaline necetigens acceptaese necetigens necetigens acceptaes acceptaese necetigens necetigens necetigens acceptaese necetigens nec

<110> 86 <110> 336

<211> STO

<213> H.Sapiens

<400> 86

Net Aso Glu Pco Leu Asp Tyr Leu Ale Aso Ala Ser Asp Phe Pro Asp 1 $^{-10}\,$

Tyr Ala Als Als Phe Gly Asn Cys Thr Asp Glu Asn Ile Pro Leo Lys $20 \ \ 20$

Met Hia Tyr Leu Pro Val Ile Tyr Gly Ile Ile Phe Leu Val Gly Phe 35 45

Pro Gly Asn Ala Val Val II: Ser Thr Tyr II: Phe Lys Met Arg Pro 50 60

Typ Lys Sor Sor The lie lie Met Leu Ash Len Ala Cys The Asp Lou 65 -76 -80

Leu Tyr Leu Thr Ser Leu Pro Phe Leu Ile His Tyr Tyr Ala Ser Gly $85 \ \ 90$

Glu Ash Trp 11e Phe Gly Asp Phe Met Cys Lys Phe 11e Arg Phe Ser 100 105 110

Fhe His Phe Asn Leu Tyr Ser Ser Ile Leu Phe Leu Thr Cys Phe Ser 115 120 125

Ile Phe Arg Tyr Cys Val Ile Ile Bis Pro Met Ser Cys Phe Ser I)e 130 135 140

His Lys Thr Arg Cys Als Val Val Als Cys Als Val Val Trp He He 145 150 155 160

Ser Leu Val Ala Val Ilo Pro Met Thr Pho Leo Ile Thr Ser Thr Ash 165 170 175

Arg Thr Ash Ary Ser Ale Cys Leu Asp Leu Thr Ser Ser Asp Glu Leu 180 185 190

Aen Thr Lie bys Trp Tyr Asn Leo Like Leo Thr Ala Ser Thr Phe Cys 195 200 205	
Leo Pro Leo Val Ile Val Thr Leo Cys Tyr Thr Thr Ilo Ule His Thr 210 215 220	
Leu Thr His Gly Leu Gln Thr Asp Ser Cys Leu Lys Gln Lys Ale Ang 235 230 235 240	
Mrq Leu Thr Ile Leu Leu Leu Leu Ala Phe Tyr Val Cys Che Leu Pro 245 250 255	
Phe Bis The Lev Arg Val. The Gln Asp Arg Ilc Ser Ala Cys Phe Gln 260 265 270	
Ser Val Val Pro Leu Arg Ile Arg Ser Met Lys Leo Thr Ser Phe Leo 275 280 285	
Asp His Tyr Ala Ata Leo Asm Thr Phe Gly Asm Leo Leo Leo Tyr Val 290 295 300	
Val. Val Sez Asp Asn Phe Glin Glin Ala Val Cys Ser Thr Val Arg Cya 305 310 315 320	
Lys Val Ser Gly Asn Leo Glo Glo Ala Lys Lys Lie Son Tyr Ser Asn 325 330 335	
<21D> 87 <211> 1138 <212> DMA <213> H.Sapiens	
<4005 87 saasattget gtaetgsact attgsatgga acttgoaaat aaagtoocht oosaaatane	60
tattetteaa cagagagtaa taggtaaatg tittagaagt gagaggacte aaal6geeaa	120
tgatttacto ttttattttt cotoctaggt ttctgggsta agtatgtgda автававаня).	180
aaadatgaga aggaadtgta adotgattat ggatttggga aaaagataaa tdaadadada	240
aagggaawag toaactgatt gwoageceto aggaatgatg contitigen acaatataat	son
taatatitee tytytymaan aenastyyte unatyatyte eytyettees tytasaytti	360
maloguechn missilongs compactor togomalous stautlatus titotatate	420
acactteasa caacttesta ecceaseaas tiggeteatt esticeatgg ecactglyga	4811
ctitettetg gggtgtetgg teatgeetta eagtatggtg agatetgetg agcastgttg	540
gtattitiggo gaagtottot gibaawattoa bacaagease gacattaige igageicage	600
ofecatitte cuttintoft Coatolocal transmictae takeoligisk gigabocact	660
gagatalaaa gogsagatga atatettegt tatttetegt atgatettea ttagttegag	720
tytocotyot ytttttyvat ttygaatyat otttotygay otaaaettoa aaggogotya Pogo 6).	780

ogagatatat	tacqaacatq	ttcactgcag	aggaggttgd	tetgtettet	ttagossaat	840
atotoggggta	otgacettta	tgacttcttt	ttatatacct	ggatetatta	tgttalgleb	900
otat@acaga	atatatetta	togothaaga	асадосьвада	ttaattagtg	atgecaatea	960
gasgotocaa	attggaliteg	aaatgaaaaa	togaatttoa	caaagcaaag	aaaqqaaaqc	1020
tgtgaagsce	ttoggoattg	tgatgggagt	thomptaata	tgetggtgcc	otttotttat	1080
ctgtacagto	atggaccott	ttottoacta	cattattoca	cotactttga	atçatyta	1138

<21**?> 88** <211> **296**

<211> 296 <212> PRT

<213> B.Sapiens

<400> 88

Met Met Pro Phe Cys His Asn Ile Ile Asn Ile Sec Cys Val Lya Asn 1 5 10 15

Asn Trp Ser Asn Asp Val Arg Ala Ser Leu Tyr Ser Leu Met Val Lou 20 25 30

He lie Leu Thr Thr Leu Val Gly Asm Leu He Val He Val Ser He $\frac{35}{45}$

Ser His Phe Lye Glo Leu Hie Thr Pro Thr Aso Top Leu The Sis Ser 50 60

Met Ala Thr Val Asp Phe Leu Leu Gly Cys Leu Val Met Pro Tyr Ser 65 70 75 80

Met Val Arg Ser Ala Glu Bis Cys Trp Tyr Phe Gly Glu Val Phe Cya 85 90 95

Lys lie His Thr Ser Thr Asp lie Met Leo Ser Ser Ala Ser lie Phe 100 105 110

His Leu Ser Phe Ile Ser Ile Asp Arg Tyr Tyr Ala Val Cys Asp Pro 115 120 125

Leu Arg Tyr Lys Alo Lys Met Asm Ile Leu Val Ile Cys Val Met Ile 130 140

Pho Ile Ser Trp Ser Val Peo Ala Vol Pho Ala Pho Gly Met Ile Pho 145 155 160

Leo Gio Leo Asn Phe Lys Gly Ala Glo Glo Ile Tyr Tyr Lys His Val $165 \,$ $170 \,$ $175 \,$

His Cys Arg Gly Gly Cys Ser val Phe Phe Ser Lys 11e Ser Gly Val

Leu Thr Phe Met Thr Sor Phe Tyr Ile Pro Gly Set Ile Met Leu Cys 195 $200\,$

Val 7γr Tyr Arg 11e Tyr Leu 11e Ala Lys Glo Glo Ala Arg Leu Ile 210 225 220 220 Ser Asp Ala Asn Glo Lys Leu Glo Ile Gly Leu Glo Met Lya Asn Gly 225 235 235 240 Ile Ser Glo Ser Lys Glo Arg Lys Ala Val Lys Thr Leu Gly Ile Val 245 250 250

Met Gly Val Phe Leu Ile Cys Tap Cys Pho Pho Pho Ile Cys Thr Val 260 265 270

Met Asp Pro Phe Leu His Tyr Ile ILe Pro Pro Thr Leo Ash Asp Ala 175 280 285

Arg Gly Ser Arg Ala Asm Ser Ala 290 - 295

<010> 89 <011> 1023

<0115 102. <0125 INA

<213> H.Sapions

<4 D02 qqaalqahqc cottttgcca caatataatt aatattteet gtgtgaassa caseluutta 60 aatgatgtee gigetteest giacagitta aiggigetea taatteigae cacasiegit 1230 180 qqcaatctga tagttattqt ttotatatca cacttcaaac aacttcatac cccaacasat tygeteatte attecutyge caetylyged thtetletgy gytytelyet eatyeettae 240 egyalgulga gatotyotga gcactgttgg tattttggag aagtettotg laasaatican 300 360 agaagracog acattatopt gageteages tecattites attiquetti cateteeatt 420 garcoctant atgotototo toatocacto agatatasao ccaagatgaa tatottyott 480 attigtgiga igaicticat tagitggagi giccotgoig tittigoatt iggsatgate 540 tttotqosqo tuaacttosa sqqoqotqaa quqatatatt acusacatot toactgoaga ngaggtiget elgheilett tagessaata telggeglad teseebbtal gaetteittt 600 660 ${\tt Lag}$ interpreted a substantial of the state of the case of the state of the s 720 caqqqaaqat taattagtga tgocaatcag aagotocaaa ttggattgga aatgaaaaat qqaatttoac aaagcaaaga aaggaaagct gtgaagacat tgqqqattgt gatgqgagtt 780 840 tipotaatat gotogteese titottiate igtaeagtea iggaesetti teiteaetae 900 attattocae otaciligas igaeghatta abibegottta gelacttasa olekacalli 960 putocaatyo thtatycatt titotatect tygittagas asgosotyas yaiqaigoty 1020 ntiggiasaa titicosaas agaitosioo aggigisaat tattitigga attgagitos

1023

<210> 90 <211> 339 <212> PRT <213> H.Sapiens

<400> 90

Met Met Pro Phe Cys Bis Asn Ile Ile Asn Ile Ser Cys Val Lys Asn 1 5 10 15

Asn Trp Ser Asn Asp Val Arg Ala Sor Lou Tyr Ser Leu Met Val Leu 20 - 30

The IIs bou The Thr Leo Val Gly Aen Leo IIe Val IIe Val Ser IIe 35 40 45

Ser His Phe Lys Glm Leu His Thr Pro Thr Asm Trp Leu Ile His Ser 50 60

Met Ala Thr Val Asp Phe Leu Leu Gly Cys Leu Val Met Pro Tyr Ser 65 70 75 80

Met Val Arg Ser Ala Clu His Cys Trp Tyr Phe Cly Glu Val Phe Cys 85 90 95

His Leu Ser Phe Ile Ser Ile Asp Arg Tyr Tyr Ala Val Cys Asp Pro 115 120 125

Leu Arg Tyr Lys Ale Lys Mot Asm Ils Loo Vel Tie Cys Vel Met Ilo 130

Fig. 11.6 Sec Trp Ser Val Pro Ala Vel Phe Ala Phe Gly Met Ile Phe 145 150 155 160

Leu Glu Leu Aan Phe Lys Gly Ala Glu Glu Tle Tyr Tyr Lys His Val 165 170 175

his Cys Arg Gly Cly Cys Ser Val Pho Pho Ser Lys Tie Ser Gly Val 180 185 180

Len The Pho Met Thr Ser Phe Tyr Ile Pro Gly Ser Ile Met Led Cys 195 200 205

Val Tyr Tyr Arg Ile Tyr Leo Ile Ala Lys Glo Gln Ala Arg Leo Ile 210 220

Ser Asp Ala Aso Glo Lya Leu Glo He Glo Wet Lys Aso Gly 225 230 235 240

The Ser Gin Ser Lys Glu Ard Lys Als Val Lys The Leo Gly Lie Val $250\,$

Met Gly Val Pho Leu lle Cys Trp Cys Pro Phe The lle Cys Thr Val Page 64

			2.60					265					270		
Met	Asp	Pro 275	Ph∈	Leju	His	тут	11 <u>e</u> 280	Il∉	Pro	Pro	This	Leb 265	Asn	Asp	Val
1%0	J.),e 290	gıŢ	Pbe	Gly	Tyr	Leu 295	Asn	Ser	Thr	Phe	Aen 300	Pro	Met	Val	Туг
Ala 305	Phe	Phe	түт	Pro	Trp 310		Arg	Lys	Ala	Leu 315	Lys	Met	Met	Leu	Phe 320
Gly	Lya	Ile	Pho	61n 325	Lys	Asp	Ser	Ser	Arrg 330	Cys	Lys	Leo	Phe	Leu 335	610
Leo	Sor	Ser													

<210> 91 <211> 1696

<212> DNA <213> B.Sapiens

<400> 91 ctytsaagta gattytatea gyactorate adgiratera etteaagtee ttegeatage 60 120 ataattacto amaxqqtqat gacaatggcy ceggyaggga tqqtgacttg cotggagatg cacageareq tolologeat acteggtest tespaceste stigstices caggeaceae 180 hocytyteca geaggastot ggggaressa aatggasast assatggaag stgasetggg 240300 tgeesetgge cacageece geneagaget tgatgatgag gactoctace eccaangtge 360 ctgggacacg gtottoctgg tggccctgct gelecttygg ctgccageca atgggttgat ggogtggotg googgotooc aggoooggoa tagaaootggo segegtotgg egotgeteef 4.30 geteageetg geneticity actiettytt cetgycagea gegycettee agateetaga galocogycat ggggggacaet ggcogotggg gacagotgdo tgcogottot actaottoot 540 ផលព្ stggggegig tectaciect coggestett estgetgque gesetempes tequesegetq ឥ៩៣ cotgotogog atgtgoccae actogtache tyggdangge pengheegen typenectets 720 gatetaneen gytytetagyy tyrtygerae actotteago gtyceotygo tygtottero cyaqqqtqqq gtqtggtggt acgaectggt catotgeetg qacttetggg acagegagga 780 840 gangingery aggangengg aggioenggg gggetteetg cettteetee toetgetegt 990 ctoccaegig ctcaeccagg ccaeaçects tequaectse cancegoesse syssessesse ageotgoogg ggottogood gtgtgoddag gaddalltelg teageotatg tegtootgag 960 1020 getgeestae ewgetgyckk agetgeteta eetggeette etgtgggaeg tetaetetgg 1000 ctacctgote tygosygees tygiciacie egactacetg atoctacies acagetyest

Page 65

cagondetto	otobapostos	tggccagtgc	egaceteegg	accetgetge	geteegtget	1140
ctogtectic	qeggeagete	totgogagga	geageeggge	agettoacge	ccactgagcc	1200
acagacecae	ctagattotg	agggtocaac	totyccagag	cogetaccas	aggueragte	1260
acagatogat	cotgtygeac	ageeteaggt	gaaccccaca	ctocagocac	gatoggabos	1320
cacageteag	ocaesgetga	accetacgge	ccagccacag	teggatecea	cagodosgod	1380
anagutņaac	ctcalingocc	agopacagto	agattetgtg	goccagocac	aggeagaeae	1440
teacqtocsy	achectocac	ctgctgccsg	ttatgligadd	agtocototg	atreagette	1500
cccaacucca	tectogeate	ctaccccagg	ggsecttgag	gacccagcca	cacotochyc	1560
ctotgaagga	gaaagcccca	gcag c acc c c	gesagaggeg	geocogggeg	caçgecese	1620
gtg-ŋgṇtsc	aggaacacgc	адарская	ņagoaņtgaa	agadoccagg	gcagacagag	1680
ульосилеся						1696

<21.0> 40

< 211 >505 FRT

<212> <213> H Sapiens

<100> 92

Leu Ala Trp Arç Oys Thr Ala Pro Ser Leu Pro Tyr Ser Val Ile His 1 10 15

The Ile He hap Ser Pro Gly The The Pro Cys Pro Ala Gly Lou Trp

Gly Pro Gln Met Asp Thr Thr Met Glu Ala Asp Leo Gly Ala Thr Gly

Ria Arg Pro Arg Thr Glu Leo Asp Asp Clu Asp Ser Tyr Pro Gln Gly

Gly Trp Asp Thr Val Phe Leu Val Ala Leu Leu Leu Leu Gly Leu Pro

Ala Asn Gly bed Met Ala Trp Led Ala Gly Ser Gln Ala Arg His Gly

Ala Gly Thr Arg Leo Ala Leo Leo Leo Ser Leo Ala Leo Ser Asp

Phe Leu Pho Lou Ala Ala Ata Ala Pho Glu Ilo Lou Glu Ilo Arg His

Gly Gly His Trp Pro Leu Gly Thr Ala Ala Cys Arg Phe Tyr Tyr Phe

Leu Trp Gly Val Ser Tyr Ser Ser Gly Leu Phe Leu Leu Ala Ala Leu 155

4	(2	Care	Lau	Less	- ו ת	Loss	Char	Pose	Ric	Trans.	Torr	Pro	C111
			Arg	165					170					175	
Eis	Arg	Pro	Va1 180	Arg	Lou	Pro	Leo	Тяр 185	Val	Сув	Ala	Gl.y	Val 190	Trp	Val
Leo	Al a	Th= 195	leu	Pb∈	Ser	Val	Pro 200	Trp	beŭ	lav	Phe	Pro 205	Glu	Als	L1.6
Val	Trp 210	Trp	Tyr	Asp	Leu	val 215	Ile	Сув	Leu	Aep	Phe 220	Trp	Аєр	Ser	Glu
Glu 225	Ləu	Ser	Leu	Arg	Met 23D	Leu	G1u	Val	Leu	Gly 235	Gly	Phe	Lec	Pro	Phe 240
Lou	Leu	Leu	Deu	Val. 245	Cys	nis	Val	Leu	Thr 250	Gln	Ala	Thr	Ala	Cys 255	yrg
Thr	Cys	ні≘	Arg 260	Gln	Gln	Glo	Pro	Ala 265	Ala	Суа	Arg	Gly	Phe 270	Ala	Arg
Val	Ala	Arg 275	Thr	Il⊖	Lou	Ser	A1@ 280	Туг	Val	val	Lou	Arg 285	Lau	Pro	Tyr
Gln	Lau 290	Ala	G1n	Len	Leu	Туп 295	Leu	$A\lambda \sigma$	Phe	Leu	Trp 300	Asp	Vaj	Ψγι	Ser
GLy 305	Tyr	Læiù	Leu	Trp	Glu 310	Ala	Leu	Val	Тук	Ser 315	Asp	Тут	Leu	Ile	Leu 32 0
Гел	Ает	Ser	Cys	Leu 325	Ser	Pro	Phe	Гел	Cys 330	Leu	Met	A.la	Ser	Ala 335	Asp
Leu	Arg	Thr	Leu 340	Leu	Arg	Ser	Val	Lew 345	Ser	Son	Phe	Ala	Ala 350	Alα	Len
Суз	Glu	Glu 355	Arg	Pro	63 y	Ser	Pho 360	The	Pro	Thr	Glu	Рло 365	Gln	Thr	Ol n
teu	Мар 370	Ser	Glu	Gly	Pro	Thr 375	Leu	Pro	Glu	Pro	Met 380	Ala	Glu	Alz	Gln
Ser 385	Gln	Met	Asp	Pro	Val 390	Ala	Gla	Pro	Gln	Val 395	Азп	Pro	Thr	Leu	G1 n 400
Pro	Arg	Ser	Авр	Pro 405	יז נליף	Mla	Glo	Pro	61n 410	Len	Asn	Pro	Thr	A18 415	GIn
PYO	Gln	ser	Аер 420	Pro	Thr	Ala	Gln	Pro 425	Gln	Leu	Asn	Leu	Met 430	Ala	Gln
Pro	Głn	Ser 435	Asp	Ser	Val	Alæ	61n 440	Pro	61 n	А)а	Азр	ፐካድ 445	Asn	Val	Gln
Thr	Pro 450	Alo	Pro	Λ1a	Ala	Ser 4S5	Ser	Vel	Pro	Ser	Pro 460	Сув	Asp	Glu	Ala
Sor	Pro	This	Pro	Ser	Ser	His	Pro	Thr		Gly Gly		Leu	Glu	Asp	Fro

465 470 475 480

Ala The Pro Pro Ala Ser Glo Gly Glo Ser Pro Ser Ser The Pro 465 490 495

Glo Ale Ale Pro Gly Ale Gly Pro Thr 500

<210> 93

<211> 1413

<212> DNA

<2130 H.Sapiena

<4D(© atygacacta nostygaago tysoctgygt gocactygod adaggnoodg cabaqagott 60 gatgatgagg actostaces commaggtggs tgggmenegg tetteetggt ggssettett 120 etrettagge tyccayecaa tyggttgatg gegtegetge competecca ggoeoggeat 160 ggagotggca egegtetage getgeteetg eteagestgg coetetetga ettettytte 240 otygoaycag oggeottees gateetsgog steepgeste gaggaesetg geogetgagg 300 acametrant gungolityka obachtocta toggyógint entactueto eggestette 360 ntquiqueny conteagest egacegetys etgetogege tytoconacs illyptanout 420 gggcscogco cagtoogcot geocototog gtotgcgcog gtgtotgggt gotggcusca 460 otottoagog tgocotogot ggtottococ gaggotgoog totggtggta cgacctygto 540 atotocoton actictopoja canegaggas ototocotoa opatoctopo 600 quettoctan etitectock polyotogta topobagiga topoccagga capagesigi 660 7.20 convections accesses as accessored a declareday appropriate total considerations. accatteigt cagodtatgi ggiootgagg eigeoctade ageiggeeda yeligebelad 730 etggeettee tgtgggaegt etactetgge tacetgetet gggaggeest ggtetastee 840 quetacelya testactosa cagetypete aquesettes tetypeteut gescugtçes 900 questorega coetgetory oterotoric technolites engagement etgenagran 960 oggeogygea getteseger esetgageea dagseeesge tagattetga yeghnewaet 1020 etgecagage egatggcaga ggeccagtea cagatggate etgtggccca geetcaggtg 1080 aaccecacae tecagocaeg ateggatese acageteage cacagetgaa eestaeggee 1140 caqueacagt equatorese agencagoes cagebysace testgeores grescartes 1200 quotototgtgg coragorada ggoagacast aacqtocaga corrtgcade tgotgoragt 1260 tetytyseea gteeetytga tyaagettee esaateeeat eetegeatee taeteeaggy 1320 queettgagg accoagodad acotectgod totgaaggang aaagdoodaag dagdacddog 13B0 ней рокосоордь ородородо аруссорае toa

1413

<210> 94 <211> 419 <212> PRT <217> 9 600

<213> н. Sapieлв

<400> 94

Most Asp Thr Thr Met Glo Ala Asp Leu Gly Ala Thr Gly Bis Arg Pro 1 $-5\,$

Arg Thr Glu Leu Aep Asp Glu Asp Ser Tyr Pro Glo Gly Gly Trp Asp 20 25 30

Thr Val Phe Leu Val Ala Leu Leu Leu Gly Leu Pro Ala Ash Gly 35 40 45

Leo Met Ala Trp Leo Ala Gly Ser Gln Ala Arg His Gly Ala Gly Thr 50 55 60

Arg Leu Als Leu Leu Leu Ser Leu Als Leu Ser Asp Rhe leu Pho 65 -70 -75 -80

Leo Ala Ala Ala Ala Phe Gin Ile Leo Glu Ile Arg His Gly Gly His 85 90 95

Trp Pro Leu Gly Thr Ala Ala Cys Arq Pho Tyr Tyr Phe Leu Trp Gly $100 \,$ $105 \,$ $110 \,$

Val Ser Tyr Ser Ser Gly Leo Phe Leo Leo Ale Ala Leo Ser Leo Asp 115 120 125

Arg Cys Leo Leo Ala Leo Cys Pro His Trp Tyr Pro Gly Dis Arg Pro 130 135: 140

Val Arg Leu Pro Leu Trp Val Cys Ala Gly Val Trp Val Leu Ala Thr 145 150 160

Len Phe Ser Val Pro Trp Len Val Phe Pro Glo Ala Val Trp Trp 165 170 175

Tyr Amp Leu Val Ile Cys Leu Amp Phe Trp Amp Ser Glu Glu Leu Sor 180 125 190

Leu Arg Met Leu Clu Val Leu Cly Cly Che Leu Pro Phe Leu Leu 195 200 205

Leu Val Cys His Val Leu Thr Gln Ala Thr Ala Cys Arg The Cys His 210 215 220

Arg Gln Gln Gln Pro Ala Alo Cys Arg Gly Phe Ala Arg Val Ala Arg 225 230 235 240

Thr fle Leu Ser Ala Tyr Val Val Leu Arg Lou Pro Tyr Cln Leu Ala 245 250 250

Glo Leo Leo Tyr Leo Ala Phe Leo Trp Asp Val Tyr Sor Gly Tyr Leo teu Trp Glo Als Leo Val Tyr Ser Asp Tyr Leo Ile Leo Leo Asm Scr Cys Leu Ser Pro Phe Leu Cys Leu Met Ala Ser Ala Asp Leu Arg Th π Leu Leu Arg Ser Val Leu Ser Ser Phe Ala Ala Ala Leu Cys Glu Glu Ang Pro Gly Sen Phe The Pro The Glo Pro Gln The Glo Leu Asp Ser Glu Gly Pro Thr Leu Pro Glu Pro Met Ala Glu Ala Glu Ser Gin Met Asp Pro Val Ala Glm Pro Glm Val Asm Pro Thr Leu Glm Pro Arg Ser Asp Pro Thr Ala Gin Pro Glo Leu Asu Pro Thr Ala Clo Pro Glo Ser Asp Pro Thr Ala Gln Pro Gln Leu Asn Leu Met Ala Gln Pro Gln Sar Asp Ser Val Ala Glm Pro Glm Ala Asp Thr Asm Val Glm Thr Pro Ala Pro Ala Ala <210> 95 <2110 49 <2120 DNA <2130 Artificial Sequence <2200 <221> misc_feature <2335 Navel Sequence <400> 95 49 ttcaasgett atggaatest etttetestt tggaglgabe ettgetgte <210> 96 49 <211> <232: DNA <2%3> Artificial Sequence <220% <221> miso feature <223> Novel Sequence <400> 96 49 tteactegag ttagecates sactetgage tggagatagt gaegatgtg Page 70

<210><211><211><212><213>	97 22 DNA Artificial Sequence	
	miss_Coature Novel Sequence	
<400> gotowar	97 ceca etcatetatg ec	22
<210 × <211 × <212 × <213 ×	96 22 DNA Artificial Sequence	
	misc feature Novel Sequence	
<400× saactte	98 etet gecettaerg te	22
<210><211><211><212><213><	99 20 DNA Artificial Sequence	
<220.5 <221.5 <223.5	Miso_feature Novel Sequence	
<400> aaagcag	99 goad beegastade	20
<210><211><211><211><212><213> 213	100 21 DNA Artificial Sequence	
<220> <221> <223>	misc_feature Novel Sequence	
<400% natgato	100 caed obgawogice d	21
<210>	101	

Page 71

<211>	27	
	DNY	
<213>	Artificial Sequence	
42127	NI III	
<220>		
	misc feature	
	Movel Sequence	
42237	Wood Todhoron	
<4000	101	
++	gett atggagtegg ggetgetg	28
E CASTA CO	dore graduation to	
<1105	316:	
< 2.118	30	
<1110 •	At(a	
c 213 ×	Artificial Sequence	
<220×		
	miso_feature	
<2230	Novel Sequence	
	• • •	
< 4(1)15	ngag tragtetgra geoggtretg	3D
T. T. C.E.J. C	Coad Pradecodra Acadacaca	
<210×	103	
:211	30	
3212		
<213/		
	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	
-2200		
< 221 ×	misc feature	
< 223 :-	Novel Sequence	
400≎	103	30
grates	stage egotatotat geocetotaeg	
48100	104	
.211		
< 0.217		
<213.5		
V = 1 = 2		
<220×		
<221>		
<223>		
<400>	104	30
cetag	agtyd acagatagog godaggalgd	
210>	105	
<211>		
<212>		
	Artificial Sequence	

PCT/U800/31581

<220>		
	misc feature	
	Novel Sequence	
	MOVEL SOMEONE	
	4.00	
<400>	105	3.0
aacccc	atca tetacaege	19
<310>	106	
	18	
	DNA	
<01.3>		
S-31, 27	Artificial Sequence	
<220>		
	misc_feature	
42235	Novel Sequence	
<400>	10.6	
		1.8
cace cal	aga geogetag	•
	104	
	33	
<212>	DMA	
<213>	Artificial Sequence	
<220>		
	misc feature	
	Nevel Sequence	
المحافقة كالمساحة	MCARI Bediance	
	405	
	107	
gdatzad	gett ceatgtacea egggtegtge tge	33
<210>	103	
	33	
	DMA	
	Artificial Sequence	
~a.1.5~	Altilicial Seguence	
<220>		
	misc feature	
< 223 >	Novel Sequence	
<4.00>	108	
	taga teagtgeeae teaacaatgt ggg	3.3
,		
-55.05	180	
	109	
	50	
	DHA	
≲213>	Artificial Segmence	
<2205		
	misc feature	

**	O 01/36473	PCT/U800/3158
<223>	Novel Sequence	
	109 mage setgtttage	20
<210><211><211><211><212><213> 4</td <td></td> <td></td>		
<220> <221> <223>	misc feature Novel Sequence	
<400° tgazat	110 acet gteegeagee	20
<2109 <2119 <2119 <2139	ONA	
<230> <231> <233>	misc feature Novel Sequence	
<4UD> gatosa	lil gott atgacaggig acticocaag taige	35
<210> <211> <210> <213>	34	
	misc_feature Novel Sequence	
<400> gateet	112 cogas gotkacoggoa caasacsosa ttoo	34
<210><210><212><212><213>	1.9 DMA	
<::20> <:221> <:223>	misc_festure	

PCT/US00/31581

Page 74

BNS(1-01) - W 0 - 16,473A2T --

PCT/US00/31581

WO 01/36473

<400> ccgcago	118 jago satgassato ag	22
	119 19 DNA Artificial Sequence	
	119 gttg tegetgase	13
<010> <011> <012> <012> <012>	120 21 DNA Artificial Sequence	
<220> <221> <223>	misc feature Novel Sequence	
<10 0 > ogatta	120 tecs caettigaco c	21
<210><211><211><211><212><213>	121 25 DNA Artiricial Sequence	
	121 catg aatgagocac tagac	25
<210><211><211><211><212><213><2213>	122 30 The The Tricial Sequence misc feature	
₹223>	Novel Sequence	
<400> ecatet	128 .cgag toaagggttg tttgagtaac	30
<210><211><211><212><213>	123 20 DNA Artificial Sequence	
<220> <221> <223>	misc feature Novel Sequence	

<400> ctgtct	123 etct guestattee	20
<210>	124	
<211>	22	
(212>	DNA	
7213>	Artificial Sequence	
	-	
<220 >		
<221:	misc feature	
	Novel Sequence	
	•	
:4000	124	
ornare day.	atet teatigaatt te	22
<2105	705	
<211>	22	
<212>	DNA	
42135	Artificial Sequence	
<220.		
	misc festure	
10 E E 200	Novel Sequence	
. 100.	105	
<4000€		22
OCCUPANT OF THE PARTY OF THE PA	aaca acttentace oo	±
. 6.4.5%	100	
	126	
<211>		
<212>		
<213>	Artificial Sequence	
<220%		
	mise feature	
3223 <i>z</i>	Novel Sequence	
	·	
<4000>	126	
acadad	agca tagtagog	19
	127	
	20	
⊴212>	DRA	
્ટો ઉં≥	Artificial Sequence	
<220≈		
<221>	misc feature	
	Novel Sequence	
	•	
<400>	127	

eagaget	tga tgatgaggac	20
<212>	128 20 DNA Artificial Sequence	
	misc_feature Novel Sequence	
<400> cocata	128 ggaa gtagtagaag	20
<210> <211> <212> <213> <223>	129 9 ERT Synthetic substrate peptide	
<221><223>	misc feature Novel Sequence	
<4000	129	
Ala Po 1	o Arg Thr ero Gly Gly Arg Arg 5	
<210><211><211><212><213>	130 52 DNA Artificial Sequence	
	mimo_feature Novel Sequence	
∴400> gegtaa	130 ataog actoactata gggaganogo etutotuota pactotutti oo	52
<210><211><211><212><213>	20 DNA	
<220> <221> <223>	misc_feature	
<400> tgcca	131 cactg atgcasetes	20

<210> <2114 <2114 <2114 <2125	132 48 DNA Artificial Sequence	
<400>	132	48
:2100	133	
<2110	24	
	DMA	
2132	Artificial Sequence	
<220>		
<221>	misc feature	
<223>	Novel Sequence	
:4000	172	
		24
<2100 <211>	134	
<211>		
<.213>		
-	,	
	124	E (3)
gogtaat	ang actemetata gggagacego augemethet timetatoco	50
42100	135	
<2115	24	
	DNA	
<21 >	Artificial Beguence	
<220>		
	misc feature	
₹223≻	Novel Sequence	
<400>	335	
	saca casticcata agec	24
	- 1, 3,	
<210> <211>	136 52	
	DATA	
(213)	Artificial Sequence	
	•	
42200		
	misc_feature Novel Sequence	
500 M	Money gedgewee	
<400>	136	52
gogtaat	taog actoactata gggagaeogo өсөмөөсмсө attocataag co	22
	Page 79	

PCT/US00/31581 WO 01/36473

<212>	137 23 ONA Artificial Sequence	
<220> <221> <223>	misc_feature Novel Sequence	
< 4 005 getacg	137 cese tetttactat coc	23
<212> <213>	138 49 DNA Artificial Sequence	
	miso feature Novel Sequence	
<400° gegtaa	138 itacq activactata Qqqaqacctt Atgaqcaqca atticaticc	49
<pre><010> <2115 <2125 <2135</pre>	20 DNA	
	misc_feature Novel Sequence	
: 400: cacae:	139 ccacc dagmaatcag	20
<2105 <2115 <2125 <2135	48 DNA	
<220× <221× <223×	misc_feature	
<400> gegta	140 ataog acteactata gggayaceca caceesceas gazatesg	48
<210	- 141 Page 80	

	21 DNA	
	Artificial Sequence	
<320°	Con Anatoma	
	misc_feature Novel Sequence	
∢ 4 (0)> ttatda	141 gcag caatteatoo c	21
	142	
:210% :212%	4KQ	
< 5.1 B ×	Artificial Sequence	
•.220 × - «: 21.»	miec feature	
	Novel Sequence	
. F. F.	140	
:4005 quqtaa	142 tagg ackgantata gggagaceeg attatonada ohbbysoon	49
<.110 · · · 111 · · ·	143	
<212 >	ANG	
<213>	Artificial Sequence	
-220:- -3205	mise feature	
	Novet Sequence	
	440	
⊲400∍ etgaaa	gtty tegetgace	19
-		
<210×		
.211> /212:		
	Artificial Sequence	
4083×	when therefore	
	misc_Ceature Novel Sequence	
<400> qcqtaa	- 144 taon actoectata gegagaccol golganagli gtogotgano	50
2.21.244		
<0.10 >		
<:11> <212>	21 PNA	
	Artificial Sections	

	misc_feature Novel Sequence	
	145 Loca cactitgado d	21
<2110 <1102		
	misc_feature Novel Sequenca	
<400> gegtaar	146 taog metuachaka gggagaceel gtaasattes cacasguace	50
<[10]> <[11]> <[12]> <[13]>	19	
<220> <321> <223>	miso_feature Novel Sequence	
៩ ៨ ៧០> គម្មាគស្នង	147 cags geaacetes	19
<211><212>	148 48 DNA Artificial Sequence	
<220><221> 20 < 20>	misc_feature Novel Sequence	
•100> ogcota	148 utae gastsastat адярациска уличасицая сванилос	48
<210> 11 1237 1237 1237	22	
<2205 <2215	misc feature	

<223>	Novel Sequence	
	149 matt cacacsagos de	22
<211> <212>	150 31 DNA Artificial Sequence	
	miac_feature Novel Sequence	
<400> geatgga	150 ntoc totttgotgt atttcaccot c	31
<111> <112>	151 31 DNA Artificial Sequence	
	miso feature Novel Sequence	
<4005 gcatgaa	151 atto acsatgodag tgstaaggaa g	31
<0112 <0182	152 31 DMA Artificial Sequence	
	misc_feature Novel Sequence	
<400> gatoaaq	152 Tott ggaatgatge cettttgeca c	31
<212>	153 29 UNA Artificial Sequence	
	misc feature Novel Seguence	

PCT/U800/31581

WO 01/36473

**	70 0120412	
<400>	153	
	ogag catcattoas agtaggtgg	2.9
GRECECE	edag caseaccas adeadarda	
<210>	154	
<211>		
	4½ AMI	
<2122 <2130		
N.S. J. 1997	Artificial Sequence	
<220>		
	misc feature	
<223>		
1.2236	MOVET REGIENCE	
<4000	154	
antart.	ogag otatgaacto asttocaasa staatttada do	4.2
darecr.	egag Geargaaces ascreenass seaserrates	
d2102	155	
72102 72112		
	49 DNA	
42115 4213>	Artificial Sequence	
*	Attititiat Sequence	
12207		
	The Feature	
	misc_feature	
6 <u>2</u> 2339	Novel Sequence	
:4000	155	
- 4000	tgaa otobacattt aatecaatgg titatgeatt titetatee	49
ger.act.	the contract of authorizing free grant for the contract of	
<210>	156	
<2115		
210		
	Artificial Sequence	
72 I C	Allifelia seguence	
<220%		
	misc leature	
4223×	Novel Sequence	
<400%	156	
onetao	aasa stgcatasac cattggatta aatgtegayt Lossephage	49
deeca	2 - 2 - 2 - 2 - 2 - 2 - 2 - 2 - 2 - 2 -	
:210	157	
211	35	
×3122	DNA	
< 21.5 >	Artificial Sequence	
< 220 -		
<221 •	misc feature	
	Novel Sequence	
<400.5	157	
gatora	satto atggacacta coatggaago tgaco	35
301040	erre cras com ext e t	

PCT/US00/31581

WO 01/36473

<210> <011> <212> <213>		
	misc feature Novel Sequence	
៤400% ដូច៤០០៤	158 ლუფი ზიოიცსფლფი იიზფილიიილ ფ	31
<2110 > <211 > <212 > <212 > <213 >	DNA	
	misc feature Novel Sequence	
<400× qugtas	149 Lacy actosotata gygagacogo gtytotypta gacbotattt oc	52
<210> <211> <212> <213>	20	
<220> <221> <023>	misc feature Novel Sequence	
<4002 tgccac	160 actg atgeameter	20
<2100 <2110 <2125 <2135	DNA	
	misc_feature Novel Sequence	
k 400 > gugtaa	161 taog gotoantala yggagaenig chanactgal gcaackee	46
<21 0 > <211>	162 24	

Page 85

<212>	DNA	
<213> <220>	Artificial Sequence	
<221>	misc_feature Novel Sequence	
****	10102 20 ,0 000	
<400> gegtgtk	162 otgo tagactotat theo	24
	163 50	
<21.22		
<220>	,	
<2210	misc feature Novel Sequence	
	•	
<400%	163 tang antoantata gggaganngo angphantot ttantaton	50
gogena	tally avecation yay y	
<210>		
<2112 <2125	DIAN	
	Artificial Sequence	
	misc_feature	
<223>	Novel Sequence	
<400 >		24
ффасал	saca ceathceata agre	
<210>		
<211> <212>	DNA	
<213>		
	misc_feature	
<223>	Novel Sequence	
<400>	165	52
gogta	atacq actemetata ggqagacege acamaacaca attecatamy ce	
<210>		
<211><211>	DNA	
<213>	Artificial Sequence	

	misc_feature Novel Sequence	
<400> gotacgo	166 pead tetttastat ped	23
<31/0 <211> <212> <213>		
	misc_feature Novel Sequence	
នុងពីបែរ ភូមិថ្នាក់ និងមិ	167 tadg adboackatи gggagaoott atgagoagoa attoatooo	19
<211> <210>	168 20 Una Artificial Sequence	
:22fb :22fb :22fb	misc_feature Novel Sequence	
	168 pacc magasatemy	20
<2112 <2122	169 48 DNA Artificial Sequence	
	misc_feature Novel Sequence	
k450s gogteat	169 Leog activoteta gggagacoca caccoaccaa gaaatcag	4 F3
<211> <212>	170 21 DNA Artificial Sequence	
<220> <221> <223>	misc_feature Novel_Sequence	

<400> ttatgag	170 geag esatteatec e	21
<210><211><211><211><012><013><	171 49 DNA Artificial Sequence	
	misc_festure Novel Sequence	
<400 × gogtaar	171 taeq acteantata upgagacess attatecaea etttgacee	49
< 212 >	172 13 DNA Artificial Sequence	
	misc feature Movel Sequence	
<400× of gasas	172 gbtq togobeacc	19
<210><210><211><212><213><	50	
<220> <221> <223>	miso_featuro Novel Sequence	
<400> gogtas	173 stacg actosotata gggagaccot gotgasagtt gtogotgacc	50
<210> <211> <212> <215>	21	
<220><221><223>	misc_feature	
< 400 %	174	

•	WO 01/36473	PCT/U800/31581
ogatio	sto ca c actttga c o c	21
<210><211>		
<212><213>	DNA	
	misc_feature Novel Sequence	
<4000 gegta:	175 stang actionista gggagaccot gtalaattos cacsagosco	50
<210> <211> <212> <2130		
	misc_foature Novel Sequence	
<400> agaaga	176 acaga gcaacotoo	19
:2100 :211> :212> :213>		
	misc_feature Novel Sequence	
<400> gogta:	177 atacg actosotata gggagaccag aagacagago aacotec	47
<010><010><010><211><212><213> 00	178 22 EMA Artificial Sequence	

Page 89

22

<220> <221> misc feature <223> Novel Sequence

ctytasaatt cecacaagca oc

 $<\!4000\times-1000$

<212:	179 31 DNA Artificial Sequence	
<220><221>	misc_feature Novel Sequence	
-4000 geat gg:	179 atoc tetttgetgt attteacest c	31
<2112 <212>	180 31 DNA Artificial Sequence	
-2200 -2210 -2230	miso_feature Novel Saquence	
.≬nQ. gcatga	ემ 0 atto acsatgocag tgataaggaa ң	31
<210 · <211 · <212 · <213 · <213 ·	181 20 DNA Artificial Sequence	
<2.20 > <2.23 > <1.23 >	mise_Feature Novel Sequence	
<400° acageo	181 pecaa agecaaacae	20
<2105 <2112 <2122 <2125	192 22 CNA Artificial Sequence	
<220> <221> <223>	miso_feature Novel Sequence	
< 4 0(0 0000.a)	182 zgago astgasaelo ag	22
<210><211><211><212>	183 20 Drw.	

9age 90

<213> Artificial Sequence	
<pre><220% <221> wisq_feature <223> Novel Sequence</pre>	
<400> 183 otgtotetet gtectattee	20
<210> 184 4311> 22 4212 DNA 4213> Artificial Sequence	
<pre><220 •</pre>	
c400: 184 genorgatet teattgactt te	22
<2105 185 (2115 1188 (2125 PNA (2137 E.Sapiens	
<400: 185 aggetegege cogaageaga geeatgagaa ceacagggtg cetggegage egetagegee	60
anggreecon quantificant goldgered tracingtha tagtactage ogtagegets	120
chatcheany canhaghest gottlythen booleanegon strangetong cactorages	180
teaggegies tranggigas teighetata ggocarcing tyringgong goiggacats	240
contrasego typetegytyt gatgegogyg neggararcyt nygonyndong ephatgedaa	300
gtoattgget beetggaeae etteetggeg teeaaegegg egetgagegt ggoggegetg	360
wipgeagage agtiggetigge agtiggette beactigget abgebiggage betigegadeg	420
ogetakgeng genbanlaget aggetafen) topagammagt ogstagenett etcapgoget	480
geacttgget götegtgget tggetaeage agugeekung ogtooligbte getgegeetg	540
pegecegage etgagegtee gegetteges geetteaceg craegetees tycogtgego	600
ttegtgetge egetggeggt getetgeete acetegetee aggtgeaeeg ggtggeaegg	660
againacture agriguativos carregioses atraagrego terregiotyet erregioty	720
caecocagig igeggeageg eigesteate mageagaage iggegengena ongogensen	780
aggaagattq gcattgctat tqcgacette etcatetget ttgccccqta tgtcatgacc	840

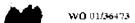
Page 91

aggetggegg agetegtged ettegtcaec gtgasegede agtggggdat cetteaguaag
tgoctgaeet acageaagge egtegoogae begttbaegt actetetget eegenggeng
ttocqueaan teetggeogg cetgetgeae oggetgetga agagaaceen gegeedagea
tecaccosts acagetetet upatelogues agentagine accageingt gaagagaace
cogegeees estecacees casegeees gbygscacks apostgatto otsportgess
cagacecset gagggcetgg cagggetest egenecesses biclassa
<2100 186 <2110 363 <2100 PRT <2100 B. Sepiens <4000 186
Met Gly Fro Gly Glu Als Leu Leu Als Gly Leo Leo Val Met Val Leu 1 5 10 15
Ala Val Ala Leu Leu Ser Asn Ala Leu Val Leu Leu Cys Cys Ala Tyr 20 25 30
Ser Ala Cio Leu Arg Thr Arg Ala Ser Gly Val Leu Leu Val Asn Leu 35
Ser Leu Gly His Leu Leu Leu Ala Ala Leu Asp Met Pro Phe Thr Leu 50 55 60
Leu Gly Val Met Arg Gly Arg Thr Pro Ser Ala Pro Gly Ala Cys Gln 65 70 75 80
Val Ila Gly Phe Lou Asp Thr Phe Leo Ala Ser Asn Ala Ala Leo Ser 85 90 95
Val Ala Ala Leo Ser Ala Asp Glo Trp Leo Ala Val Gly Pho Pro Leo 100 105 110
Arg Tyr Ala Gly Arg Leu Arg Pro Arg Tyr Ala Gly Leu Leu Leu Gly 115 120 125
Cyp Ala Trp Cly Gln Ser Leu Ala Phe Ser Gly Ala Ala Leo Gly Cys 130 135 140
Sor Trp Lou Gly Tyr Ser Ser Ale Phe Ala Ser Cys Ser Leu Arg Leu 145 150 155 160
Pro Pro Glu Pro Glu Arg Pro Arg Phe Ale Ale Phe Thr Ale The Lev 165 170 175
Ris Ala Val Gly Phe Val Leu Pro Leu Ala Val Leu Cya Leu Thr Ser 180 195 190
Leu Gln Mal His Arg Wal Ala Arg Arg Dis Cys Gln Arg Met Asp Thr 195 200 205



29

28



۷al	200 Thr	Mot	Ly5	BIA	Leo	Ala 215	Leu	Len	είπ	Asp	Leu 220	иis	Pro	Ser	Val
Arç 225	GIV	Arg	Cys	Гел	11e 230	Gln	Gln	Ьу≲	Arç	Arg 235	Arg	ніз	Arg	Ala	Thr 240
Arg	L∙ya	lle	Gly	Ile 245	Ala	He	Ala	Thr	Phe 250	Leu	Ile	Cys	Phe	Ala 255	Pro
Туг	Val	Met	Thr 260	Arg	Leo	Ala	Glu	Leu 265	Val	Pro	Ьре	Val	Thr 270	Val	Asn
Ala	Gln	Trp 275	СГА	Ile	Leo	Ser	Lys 280	Cys	Lep	Thr	Тус	5⊕x 265	Lys	Λla	Vel
ħΙ.a	Asp 290	Ρτο	Phe	Thr	Tyr	Ser 2 9 5	Leu	Lev	Arg	Arg	Pro 300	Fhe	Arg	Gln	Vel
Leu 3D5	Ala	Gly	Met	Val	Sis 310	Arg	Lev	Leu	Lys	Arŋ 315	Thr	Pro	Arg	Fre	Ala 320
Ser	Thr	Bis	Азр	Ser 325	Ser	Long	Asp	Val	Ala 330	Gλу	Mot	Val	His	61n 335	Leu
Lau	Lys	Arg	The 340	Pro	Arg	Pro	Als	Ser 345	Thr	ніє	Aen	Gly	Ser 350	Val	Asp
Thr	Glu	88b 355	Аар	Ser	Cys	Leu	G1n 360	Gln	Thr	His					
<210		187													
<211 <212		29 DNA													
<2135 Artificial Sequence															
<226	<22(0>														
<221> misc_feature <223> Novel Sequence															
, Z Z Z	(i,2	MOARI	r ser	Inenc	: 0									•	
<400	ja	187													
gcataagett gecatygges coggegagg															
		2 (5.13													
<210 <211		100 28													
<212		DNA		5 2.											
<213		Artii	11013	:1 56	dnet	100									
<220 <2 2 1		miac	fest	1150											
<220		Novel			e										

<210> 189

<400> 1.8H

qualitotaga cotcagtgtg totgotgo

Page 93

WO 01/36473

PCT/US00/31581

<311>	50	
<212>		
<213>	Artificial Sequence	
<220>		
	misc feature	
<223>	Novel Sequence	
-:400>	109	
tigatiqa	ttte tteograntes	20
	A.C.D.	
<210>		
<211>		
<212>		
<213>	Artificial Sequence	
<230>		
<cc1:< td=""><td>máso_feature</td><td></td></cc1:<>	máso_feature	
00233°	Newol Sequence	
<40105	250	
titquac	gera ggaaggtg	18